CIRCUIT ANALYSIS and FEEDBACK AMPLIFIER THEORY

CIRCUIT ANALYSIS and FEEDBACK AMPLIFIER THEORY

Edited by
Wai-Kai Chen
University of Illinois
Chicago, U.S.A.

Taylor & Francis
Taylor & Francis Group
Boca Raton London New York

A CRC title, part of the Taylor & Francis imprint, a member of the
Taylor & Francis Group, the academic division of T&F Informa plc.

Published in 2006 by
CRC Press
Taylor & Francis Group
6000 Broken Sound Parkway NW, Suite 300
Boca Raton, FL 33487-2742

Printed in the United States of America on acid-free paper
10 9 8 7 6 5 4 3 2 1

International Standard Book Number-10: 0-8493-5699-7 (Hardcover)
International Standard Book Number-13: 978-0-8493-5699-5 (Hardcover)

Library of Congress Cataloging-in-Publication Data

Catalog record is available from the Library of Congress

T&F informa
Taylor & Francis Group
is the Academic Division of T&F Informa plc.

Visit the Taylor & Francis Web site at
http://www.taylorandfrancis.com

and the CRC Press Web site at
http://www.crcpress.com

Preface

The purpose of Circuit Analysis and Feedback Amplifier Theory is to provide in a single volume a comprehensive reference work covering the broad spectrum of linear circuit analysis and feedback amplifier design. It also includes the design of multiple-loop feedback amplifiers. The book is written and developed for the practicing electrical engineers in industry, government, and academia. The goal is to provide the most up-to-date information in the field.

Over the years, the fundamentals of the field have evolved to include a wide range of topics and a broad range of practice. To encompass such a wide range of knowledge, the book focuses on the key concepts, models, and equations that enable the design engineer to analyze, design and predict the behavior of large-scale circuits and feedback amplifiers. While design formulas and tables are listed, emphasis is placed on the key concepts and theories underlying the processes.

The book stresses fundamental theory behind professional applications. In order to do so, it is reinforced with frequent examples. Extensive development of theory and details of proofs have been omitted. The reader is assumed to have a certain degree of sophistication and experience. However, brief reviews of theories, principles and mathematics of some subject areas are given. These reviews have been done concisely with perception.

The compilation of this book would not have been possible without the dedication and efforts of Professor Larry P. Huelsman, and most of all the contributing authors. I wish to thank them all.

Wai-Kai Chen
Editor-in-Chief

Editor-in-Chief

Wai-Kai Chen, Professor and Head Emeritus of the Department of Electrical Engineering and Computer Science at the University of Illinois at Chicago, is now serving as Academic Vice President at International Technological University. He received his B.S. and M.S. degrees in electrical engineering at Ohio University, where he was later recognized as a Distinguished Professor. He earned his Ph.D. in electrical engineering at the University of Illinois at Urbana/Champaign.

Professor Chen has extensive experience in education and industry and is very active professionally in the fields of circuits and systems. He has served as visiting professor at Purdue University, University of Hawaii at Manoa, and Chuo University in Tokyo, Japan. He was Editor of the *IEEE Transactions on Circuits and Systems*, *Series I and II*, President of the IEEE Circuits and Systems Society, and is the Founding Editor and Editor-in-Chief of the *Journal of Circuits, Systems and Computers*. He received the Lester R. Ford Award from the Mathematical Association of America, the Alexander von Humboldt Award from Germany, the JSPS Fellowship Award from Japan Society for the Promotion of Science, the Ohio University Alumni Medal of Merit for Distinguished Achievement in Engineering Education, the Senior University Scholar Award and the 2000 Faculty Research Award from the University of Illinois at Chicago, and the Distinguished Alumnus Award from the University of Illinois at Urbana/Champaign. He is the recipient of the Golden Jubilee Medal, the Education Award, the Meritorious Service Award from IEEE Circuits and Systems Society, and the Third Millennium Medal from the IEEE. He has also received more than a dozen honorary professorship awards from major institutions in China.

A fellow of the Institute of Electrical and Electronics Engineers and the American Association for the Advancement of Science, Professor Chen is widely known in the profession for his *Applied Graph Theory* (North-Holland), *Theory and Design of Broadband Matching Networks* (Pergamon Press), *Active Network and Feedback Amplifier Theory* (McGraw-Hill), *Linear Networks and Systems* (Brooks/Cole), *Passive and Active Filters: Theory and Implements* (John Wiley), *Theory of Nets: Flows in Networks* (Wiley-Interscience), and *The VLSI Handbook* (CRC Press).

Advisory Board

Contributors

Peter Aronhime
University of Louisville
Louisville, Kentucky

K.S. Chao
Texas Tech University
Lubbock, Texas

Ray R. Chen
San Jose State University
San Jose, California

Wai-Kai Chen
University of Illinois
Chicago, Illinois

John Choma, Jr.
University of Southern California
Los Angeles, California

Artice M. Davis
San Jose State University
San Jose, California

Marwan M. Hassoun
Iowa State University
Ames, Iowa

Pen-Min Lin
Purdue University
West Lafayette, Indiana

Robert W. Newcomb
University of Maryland
College Park, Maryland

Benedykt S. Rodanski
University of Technology, Sydney
Broadway, New South Wales,
Australia

Marwan A. Simaan
University of Pittsburgh
Pittsburgh, Pennsylvania

James A. Svoboda
Clarkson University
Potsdam, New York

Jiri Vlach
University of Waterloo
Waterloo, Ontario, Canada

Table of Contents

1

Fundamental Circuit Concepts

John Choma, Jr.
University of Southern California

1.1 The Electrical Circuit

An **electrical circuit** or **electrical network** is an array of interconnected elements wired so as to be capable of conducting current. As discussed earlier, the fundamental **two-terminal elements** of an electrical circuit are the **resistor**, the **capacitor**, the **inductor**, the **voltage source**, and the **current source**. The circuit schematic symbols of these elements, together with the algebraic symbols used to denote their respective general values, appear in Figure 1.1.

As suggested in Figure 1.1, the value of a resistor is known as its *resistance*, R, and its dimensional units are *ohms*. The case of a wire used to interconnect the terminals of two electrical elements corresponds to the special case of a resistor whose resistance is ideally zero ohms; that is, $R = 0$. For the capacitor in Figure 1.1(b), the *capacitance*, C, has units of *farads,* and from Figure 1.1(c), the value of an inductor is its *inductance*, L, the dimensions of which are *henries.* In the case of the voltage sources depicted in Figure 1.1(d), a constant, time invariant source of voltage, or *battery*, is distinguished from a voltage source that varies with time. The latter type of voltage source is often referred to as a **time varying signal** or simply, a **signal**. In either case, the value of the battery voltage, E, and the time varying signal, $v(t)$, is in units of *volts*. Finally, the current source of Figure 1.1(e) has a value, I, in units of *amperes*, which is typically abbreviated as amps.

Elements having three, four, or more than four terminals can also appear in practical electrical networks. The discrete component **bipolar junction transistor** (BJT), which is schematically portrayed in Figure 1.2(a), is an example of a three-terminal element, in which the three terminals are the collector, the base, and the emitter. On the other hand, the monolithic *metal-oxide-semiconductor field-effect transistor* (MOSFET) depicted in Figure 1.2(b) has four terminals: the drain, the gate, the source, and the bulk substrate.

Multiterminal elements appearing in circuits identified for systematic mathematical analyses are routinely represented, or *modeled*, by equivalent subcircuits formed of only interconnected two-terminal elements. Such a representation is always possible, provided that the list of two-terminal elements itemized in Figure 1.1 is appended by an additional type of two-terminal element known as the **controlled source**, or **dependent generator**. Two of the four types of controlled sources are voltage sources and two are current sources. In Figure 1.3(a), the dependent generator is a *voltage-controlled voltage source* (VCVS) in that the voltage, $v_0(t)$, developed from terminal 3 to terminal 4 is a function of, and is therefore

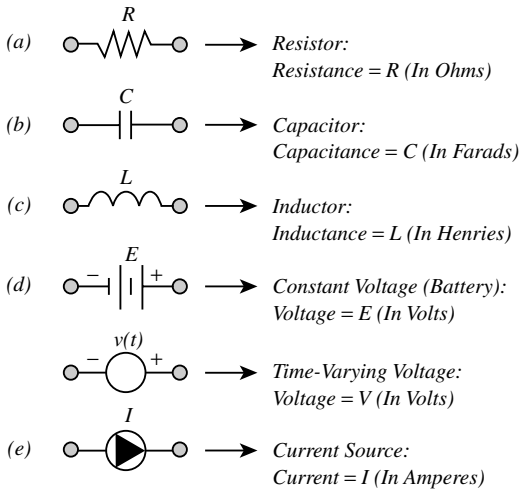


FIGURE 1.1 Circuit schematic symbol and corresponding value notation for (a) resistor, (b) capacitor, (c) inductor, (d) voltage source, and (e) current source. Note that a constant voltage source, or battery, is distinguished from a voltage source that varies with time.

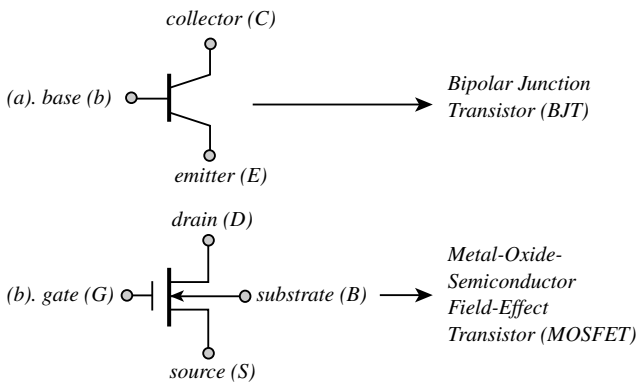


FIGURE 1.2 Circuit schematic symbol for (a) discrete component bipolar junction transistor (BJT) and (b) monolithic metal-oxide-semiconductor field-effect transistor (MOSFET).

dependent on, the voltage, $v_i(t)$, established elsewhere in the considered network from terminal 1 to terminal 2. The *controlled voltage*, $v_0(t)$, as well as the *controlling voltage*, $v_i(t)$, can be constant or time varying. Regardless of the time-domain nature of these two voltage, the value of $v_0(t)$ is not an independent number. Instead, its value is determined by $v_i(t)$ in accordance with a prescribed functional relationship, e.g.,

$$v_0(t) = f\left[v_i(t)\right] \tag{1.1}$$

If the function, $f(\cdot)$, is linearly related to its argument, (1.1) collapses to the form

$$v_0(t) = f_\mu v_i(t) \tag{1.2}$$

where $f\mu$ is a constant, independent of either $v_0(t)$ or $v_i(t)$. When the function on the right-hand side of (1.1) is linear, the subject VCVS becomes known as a *linear voltage-controlled voltage source.*

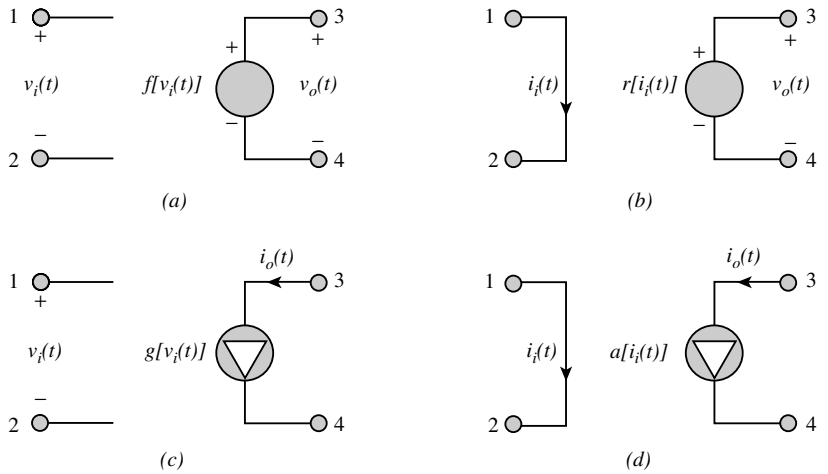


FIGURE 1.3 Circuit schematic symbol for (a) voltage-controlled voltage source (VCVS), (b) current-controlled voltage source (CCVS), (c) voltage-controlled current source (VCCS), and (d) current-controlled current source (CCCS).

The second type of controlled voltage source is the *current-controlled voltage source* (CCVS) depicted in Figure 1.3(b). In this dependent generator, the controlled voltage, $v_0(t)$, developed from terminal 3 to terminal 4 is a function of the *controlling current*, $i_i(t)$, flowing elsewhere in the network between terminals 1 and 2, as indicated. In this case, the generalized functional dependence of $v_0(t)$ on $i_i(t)$ is expressible as

$$v_0(t) = r\left[i_i(t)\right] \tag{1.3}$$

which reduces to

$$v_0(t) = r_m i_i(t) \tag{1.4}$$

when r(·) is a linear function of its argument.

The two types of dependent current sources are diagrammed symbolically in Figures 1.3(c) and (d). Figure 1.3(c) depicts a *voltage-controlled current source* (VCCS), for which the controlled current $i_0(t)$, flowing in the electrical path from terminal 3 to terminal 4, is determined by the controlling voltage, $v_i(t)$, established across terminals 1 and 2. Therefore, the controlled current can be written as

$$i_0(t) = g\left[v_i(t)\right] \tag{1.5}$$

In the *current-controlled current source* (CCCS) of Figure 1.3(d),

$$i_0(t) = a\left[i_i(t)\right] \tag{1.6}$$

where the controlled current, $i_0(t)$, flowing from terminal 3 to terminal 4 is a function of the controlling current, $i_i(t)$, flowing elsewhere in the circuit from terminal 1 to terminal 2. As is the case with the two controlled voltage sources studied earlier, the preceding two equations collapse to the linear relationships

$$i_0(t) = g_m v_i(t) \tag{1.7}$$

and

$$i_0(t) = a_\alpha i_i(t) \tag{1.8}$$

when $g(\cdot)$ and $a(\cdot)$, respectively, are linear functions of their arguments.

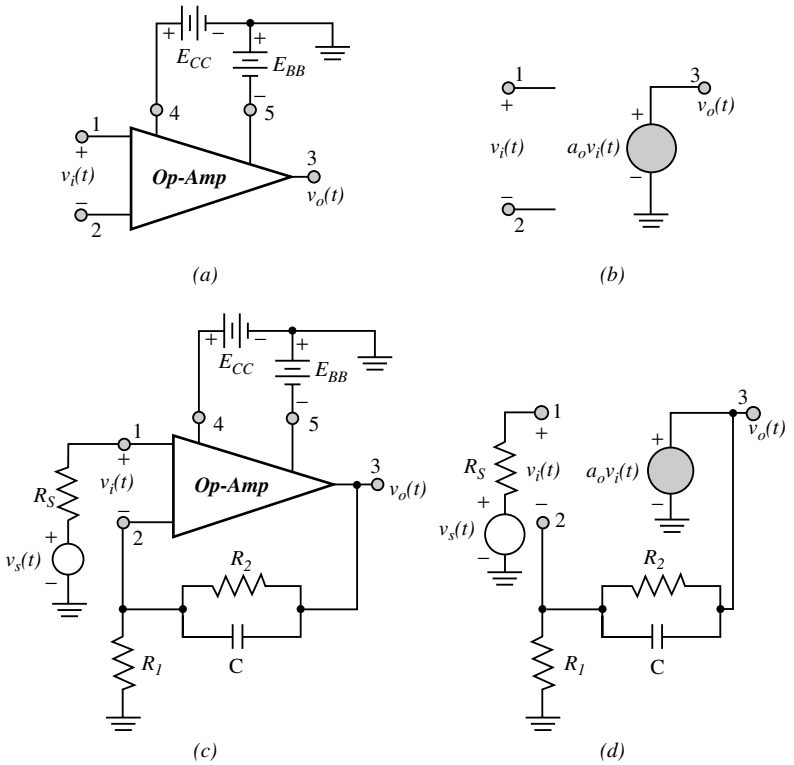


FIGURE 1.4 (a) Circuit schematic symbol for a voltage mode operational amplifier. (b) First-order linear model of the op-amp. (c) A voltage amplifier realized with the op-amp functioning as the gain element. (d) Equivalent circuit of the voltage amplifier in (c).

The immediate implication of the controlled source concept is that the definition for an electrical circuit given at the beginning of this subsection can be revised to read "an electrical circuit or electrical network is an array of *interconnected two-terminal elements* wired in such a way as to be capable of conducting current". Implicit in this revised definition is the understanding that the two-terminal elements allowed in an electrical circuit are the resistor, the capacitor, the inductor, the voltage source, the current source, and any of the four possible types of dependent generators.

In, an attempt to reinforce the engineering utility of the foregoing definition, consider the voltage mode **operational amplifier**, or **op-amp**, whose circuit schematic symbol is submitted in Figure 1.4(a). Observe that the op-amp is a five-terminal element. Two terminals, labeled 1 and 2, are provided to receive input signals that derive either from external signal sources or from the output terminals of subcircuits that feed back a designable fraction of the output signal established between terminal 3 and the system ground. Battery voltages, identified as E_{CC} and E_{BB} in the figure, are applied to the remaining two op-amp terminals (terminals 4 and 5) with respect to ground to *bias* or *activate* the op-amp for its intended application. When E_{CC} and E_{BB} are selected to ensure that the subject op-amp behaves as a linear circuit element, the voltages, E_{CC} and E_{BB}, along with the corresponding terminals at which they are incident, are inconsequential. In this event the op-amp of Figure 1.4(a) can be modeled by the electrical circuit appearing in Figure 1.4(b), which exploits a linear VCVS. Thus, the voltage amplifier of Figure 1.4(c), which interconnects two batteries, a signal source voltage, three resistors, a capacitor, and an op-amp, can be represented by the network given in Figure 1.4(d). Note that the latter configuration uses only two terminal elements, one of which is a VCVS.

Current and Current Polarity

The concept of an electrical *current* is implicit to the definition of an electrical circuit in that a circuit is said to be an array of two-terminal elements that are connected in such a way as to permit the *condition of current*. Current flow through an element that is capable of current conduction requires that the net *charge* observed at any elemental cross-section change with time. Equivalently, a net nonzero charge, $q(t)$, must be transferred over finite time across any cross-sectional area of the element. The current, $i(t)$, that actually flows is the time rate of change of this transferred charge;

$$i(t) = \frac{dq(t)}{dt} \tag{1.9}$$

where the MKS unit of charge is the *coulomb*, time t is measured in seconds, and the resultant current is measured in units of amperes. Note that zero current does not necessarily imply a lack of charge at a given cross-section of a conductive element. Instead, zero current implies only that the subject charge is not changing with time; that is, the charge is not moving through the elemental cross-section.

Electrical charge can be negative, as in the case of *electrons* transported through a cross-section of a conductive element such as aluminum or copper. A single electron has a charge of $-(1.6021 \times 10^{-19})$ coulomb. Thus, (1.9) implies a need to transport an average of (6.242×10^{18}) electrons in 1 second through a cross-section of aluminum if the aluminum element is to conduct a constant current of 1 amp. Charge can also be positive, as in the case of *holes* transported through a cross-section of a semiconductor such as germanium or silicon. Hole transport in a semiconductor is actually electron transport at an energy level that is smaller than the energy required to effect electron transport in that semiconductor. To first order, therefore, the electrical charge of a hole is the negative of the charge of an electron, which implies that the charge of a hole is $+(1.602 \times 10^{-19})$ coulomb.

A positive charge, $q(t)$, transported from the left of the cross-section to the right of the cross-section in the element abstracted in Figure 1.5(a) gives rise to a positive current, $i(t)$, which also flows from left to right across the indicated cross-section. Assume that, prior to the transport of such charge, the volumes to the left and to the right of the cross-section are electrically neutral; that is, these volumes have zero initial net charge. Then, the transport of a positive charge, q_0, from the left side to the right side of the element charges the right side to $+1q_0$ and the left side to $-1q_0$.

Alternatively, suppose a negative charge in the amount of $-q_0$ is transported from the right side of the element to its left side, as suggested in Figure 1.5(b). Then, the left side charges to $-q_0$, and the right side charges to $+q_0$, which is identical to the electrostatic condition incurred by the transport of a positive charge in the amount of q_0 from left- to right-hand sides. As a result, the transport of a net negative charge from right to left produces a positive current, $i(t)$, flowing from left to right, just as positive charge transported from left- to right-hand sides induces a current flow from left to right.

Assume, as portrayed in Figure 1.5(c), that a positive or a negative charge, say, $q_1(t)$, is transported from the left side of the indicated cross-section to the right side. Simultaneously, a positive or a negative charge in the amount of $q_2(t)$ is directed through the cross-section from right to left. If $i_1(t)$ is the current arising from the transport of the charge $q_1(t)$, and if $i_2(t)$ denotes the current corresponding to the transport of the charge, $q_2(t)$, the net effective current $i_e(t)$, flowing from the left side of the cross-section to the right side of the cross-section is

$$i_e(t) = \frac{d}{dt}\left[q_1(t) - q_2(t)\right] = i_1(t) - i_2(t) \tag{1.10}$$

where the charge difference, $[q_1(t) - q_2(t)]$, represents the net charge transported from left to right. Observe that if $q_1(t) \equiv q_2(t)$, the net effective current is zero, even though conceivably large numbers of charges are transported back and forth across the junction.

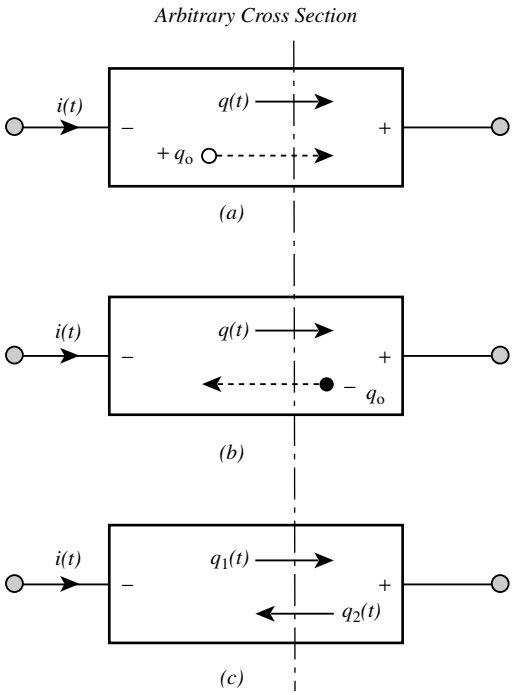


FIGURE 1.5 (a) Transport of a positive charge from the left-hand side to the right-hand side of an arbitrary cross-section of a conductive element. (b) Transport of a negative charge from the right-hand side to the left-hand side of an arbitrary cross-section of a conductive element. (c) Transport of positive or negative charges from either side to the other side of an arbitrary cross-section of a conductive element.

Energy and Voltage

The preceding section highlights the fundamental physical fact that the flow of current through a conductive electrical element mandates that a net charge be transported over finite time across any arbitrary cross-section of that element. The electrical effect of this charge transport is a net positive charge induced on one side of the element in question and a net negative charge (equal in magnitude to the aforementioned positive charge) mirrored on the other side of the element. This ramification conflicts with the observable electrical properties of an element in equilibrium. In particular, an element sitting in free space, without any electrical connection to a source of *energy*, is necessarily in equilibrium in the sense that the net positive charge in any volume of the element is precisely counteracted by an equal amount of charge of opposite sign in said volume. Thus, if none of the elements abstracted in Figure 1.5 is connected to an external source of energy, it is physically impossible to achieve the indicated electrical charge differential that materializes across an arbitrary cross-section of the element when charge is transferred from one side of the cross-section to the other.

The energy commensurate with sustaining current flow through an electrical element derives from the application of a *voltage*, $v(t)$, across the element in question. Equivalently, the application of electrical energy to an element manifests itself as a voltage developed across the terminals of an element to which energy is supplied. The amount of applied voltage, $v(t)$, required to sustain the flow of current, $i(t)$, as diagrammed in Figure 1.6(a), is precisely the voltage required to offset the electrostatic implications of the differential charge induced across the element through which $i(t)$ flows. This is to say that without the connection of the voltage, $v(t)$, to the element in Figure 1.6(a), the element cannot be in equilibrium. With $v(t)$ connected, equilibrium for the entire system comprised of element and voltage source is reestablished by allowing for the conduction of the current, $i(t)$.

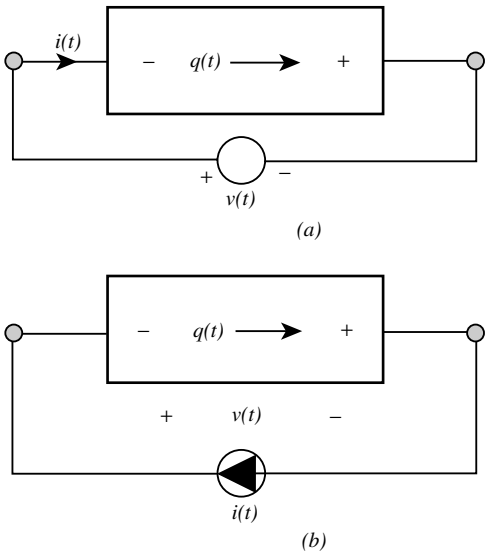


FIGURE 1.6 (a) The application of energy in the form of a voltage applied to an element that is made to conduct a specified current. The applied voltage, $v(t)$, causes the current, $i(t)$, to flow. (b) The application of energy in the form of a current applied to an element that is made to establish a specified terminal voltage. The applied current, $i(t)$, causes the voltage, $v(t)$, to be developed across the terminals of the electrical element.

Instead of viewing the delivery of energy to an electrical element as the ramification of a voltage source applied to the element, the energy delivery may be interpreted as the upshot of a current source used to excite the element, as depicted in Figure 1.6(b). This interpretation follows from the fact that energy must be applied in an amount that effects charge transport at a desired time rate of change. It follows that the application of a current source in the amount of the desired current is necessarily in one-to-one correspondence with the voltage required to offset the charge differential manifested by the charge transport that yields the subject current. To be sure, a voltage source is a physical entity, while current source is not; but the mathematical modeling of energy delivery to an electrical element can nonetheless be accomplished through either a voltage source or a current source.

In Figure 1.6, the terminal voltage, $v(t)$, corresponding to the energy, $w(t)$, required to transfer an amount of charge, $q(t)$, across an arbitrary cross-section of the element is

$$v(t) = \frac{dw(t)}{dq(t)} \tag{1.11}$$

where $v(t)$ is in units of volts when $q(t)$ is expressed in coulombs, and $w(t)$ is specified in joules. Thus, if 1 joule of applied energy results in the transport of 1 coulomb of charge through an element, the elemental terminal voltage manifested by the 1 joule of applied energy is 1 volt.

It should be understood that the derivative on the right-hand side of (1.11), and thus the terminal voltage demanded of an element that is transporting a certain amount of charge through its cross-section, is a function of the properties of the type of material from which the element undergoing study is fabricated. For example, an insulator such as paper, air, or silicon dioxide is ideally incapable of current conduction and hence, intrinsic charge transport. Thus, $q(t)$ is essentially zero in an insulator and by (1.11), an infinitely large terminal voltage is required for even the smallest possible current. In a conductor such as aluminum, iron, or copper, large amounts of charge can be transported for very small applied energies. Accordingly, the requisite terminal voltage for even very large currents approaches zero in ideal conductors. The electrical properties of semiconductors such as germanium, silicon, and gallium arsenide

lie between the extremes of those for an insulator and a conductor. In particular, semiconductor elements behave as insulators when their terminals are subjected to small voltages, while progressively larger terminal voltages render the electrical behavior of semiconductors akin to conductors. This conditional conductive property of a semiconductor explains why semiconductor devices and circuits generally must be biased to appropriate voltage levels before these devices and circuits can function in accordance with their requirements.

Power

The foregoing material underscores the fact that the flow of current through a two-terminal element, or more generally, through any two terminals of an electrical network, requires that charge be transported over time across any cross-section of that element or network. In turn, such charge transport requires that energy be supplied to the network, usually through the application of an external voltage source. The time rate of change of this applied energy is the *power* delivered by the external voltage or current source to the network in question. If $p(t)$ denotes this power in units of watts

$$p(t) = \frac{dw(t)}{dq} \tag{1.12}$$

where, of course, $w(t)$ is the energy supplied to the network in joules. By rewriting (1.12) in the form

$$p(t) = \frac{dw(t)}{dq(t)} \frac{dq(t)}{dt} \tag{1.13}$$

and applying (1.9) and (1.11), the power supplied to the two terminals of an element or a network becomes the more expedient relationship

$$p(t) = v(t)i(t) \tag{1.14}$$

Equation (1.14) expresses the power delivered to an element as a simple product of the voltage applied across the terminals of the element and the resultant current conducted by that element. However, care must be exercised with respect to relative voltage and current polarity, when applying (1.14) to practical circuits.

To the foregoing end, it is useful to revisit the simple abstraction of Figure 1.6(a), which is redrawn as the slightly modified form in Figure 1.7. In this circuit, a signal source voltage, $v_s(t)$, is applied across the two terminals, 1 and 2, of an element, which responds by conducting a current $i(t)$, from terminal 1 to terminal 2 and developing a corresponding terminal voltage $v(t)$, as illustrated. If the wires (zero resistance conductors, as might be approximated by either aluminum or copper interconnects) that connect the signal source to the element are ideal, the voltage, $v(t)$, is identical to $v_s(t)$. Moreover, because the current is manifested by the application of the signal source, which thereby establishes a closed electrical path for current conduction, the element current, $i(t)$, is necessarily the same as the current, $i_s(t)$, that flows through $v_s(t)$.

If attention is focused on only the element in Figure 1.7, it is natural to presume that the current conducted by the element actually flows from terminal 1 to terminal 2 when (as shown) the voltage developed across the element is positive at terminal 1 with respect to terminal 2. This assertion may be rationalized qualitatively by noting that the positive voltage nature at terminal 1 acts to repel positive charges from terminal 1 to terminal 2, where the negative nature of the developed voltage, $v(t)$, tends to attract the repulsed positive charges. Similarly, the positive nature of the voltage at terminal 1 serves to attract negative charges from terminal 2, where the negative nature of $v(t)$ tends to repel such negative charges. Because current flows in the direction of transported positive charge and opposite to the direction of transported negative charge, either interpretation gives rise to an elemental current, $i(t)$, which flows

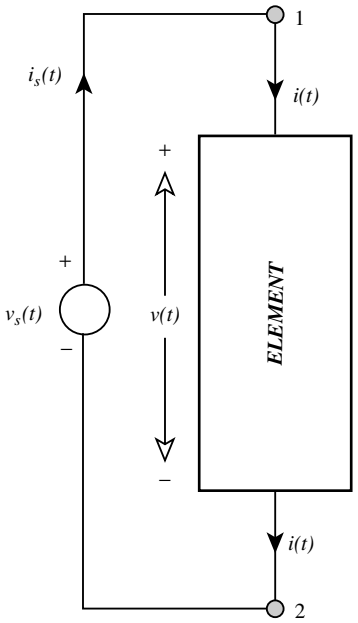


FIGURE 1.7 Circuit used to illustrate power calculations and the associated reference polarity convention.

from terminal 1 to terminal 2. In general, if current is indicated as flowing from the "high" (+) voltage terminal to the "low" (–) voltage terminal of an element, the current conducted by the element and the voltage developed across the element to cause this flow of current are said to be in **associated reference polarity**. When the element current, $i(t)$, and the corresponding element voltage, $v(t)$, as exploited in the defining power relationship of (1.14), are in associated reference polarity, the resulting computed power is a positive number and is said to represent the *power delivered to* the element. In contrast, $v(t)$ and $i(t)$ are said to be in **disassociated reference polarity** when $i(t)$ flows from the "low" voltage terminal of the element to its "high" voltage terminal. In this case the voltage-current product in (1.14) is a negative number. Instead of stating that the resulting negative power is delivered to the element, it is more meaningful to assert that the computed negative power is a *positive power that is generated by* the element in question.

At first glance, it may appear as though the latter polarity disassociation between element voltage and current variables is an impossible circumstance. Not only is polarity disassociation possible, it is absolutely necessary if electrical circuits are to subscribe to the fundamental principle of **conservation of power**. This principle states that the net power dissipated by a circuit must be identical to the net power supplied to that circuit. A confirmation of this basic principle derives from a further consideration of the topology in Figure 1.7. The electrical variables, $v(t)$ and $i(t)$, pertinent to the element delineated in this circuit, are in associated reference polarity. Accordingly, the power, $p_e(t)$, dissipated by this element is positive and given by (1.14):

$$p_e(t) = v(t)i(t) \tag{1.15}$$

However, the voltage and current variables, $v_s(t)$ and $i_s(t)$, relative to the signal source voltage are in disassociated polarity. It follows that the power, $p_s(t)$, *delivered to the signal source* is

$$p_s(t) = -v_s(t)i_s(t) \tag{1.16}$$

Because, as stated previously, $v_s(t) = v(t)$ and $i_s(t) = i(t)$, for the circuit at hand, (1.16) can be written as

$$p_s(t) = -v(t)i(t) \tag{1.17}$$

The last result implies that the

$$\text{power } \textbf{delivered by} \text{ the signal source} = +v(t)i(t) \equiv p_e(t) \tag{1.18}$$

that is, the power delivered to the element by the signal source is equal to the power dissipated by the element.

An alternative statement to conservation of power, as applied to the circuit in Figure 1.7 derives from combining (1.15) and (1.17) to arrive at

$$p_s(t) + p_e(t) = 0 \tag{1.19}$$

The foregoing result may be generalized to the case of a more complex circuit comprised of an electrical interconnection of N elements, some of which may be voltage and current sources. Let the voltage across the kth element by $v_k(t)$, and let the current flowing through this kth element, in associated reference polarity with $v_k(t)$, be $i_k(t)$. Then, the power, $p_k(t)$, delivered to the kth electrical element is $v_k(t)\ i_k(t)$. By conservation of power,

$$\sum_{k=1}^{N} p_k(t) = \sum_{k=1}^{N} v_k(t) = i_k(t) = 0 \tag{1.20}$$

The satisfaction of the expression requires that at least one of the $p_k(t)$ be negative, or equivalently, at least one of the N elements embedded in the circuit at hand must be a source of energy.

1.2 Circuit Classifications

It was pointed out earlier that the relationship between the current that is made to flow through an electrical element and the applied energy, and thus voltage, that is required to sustain such current flow is dictated by the material from which the subject element is fabricated. The element material and the associated manufacturing methods exploited to realize a particular type of circuit element determine the mathematical nature between the voltage applied across the terminals of the element and the resultant current flowing through the element. To this end, electrical elements and circuits in which they are embedded are generally codified as linear or nonlinear, active or passive, time varying or time invariant, and lumped or distributed.

Linear vs. Nonlinear

A **linear two-terminal circuit element** is one for which the voltage developed across, and the current flowing through, are related to one another by a linear algebraic or a linear integro-differential equation. If the relationship between terminal voltage and corresponding current is nonlinear, the element is said to be *nonlinear*. A linear circuit contains only linear circuit elements, while a circuit is said to be nonlinear if a least one of its embedded electrical elements is nonlinear.

All practical circuit elements, and thus all practical electrical networks, are inherently nonlinear. However, over suitably restricted ranges of applied voltages and corresponding currents, the volt-ampere characteristics of these elements and networks emulate idealized linear relationships. In the design of an electronic linear signal processor, such as an amplifier, an implicit engineering task is the implementation of biasing subcircuits that constrain the voltages and currents of internal semiconductor elements to ranges that ensure linear elemental behavior over all possible operating conditions.

The voltage–current relationship for the linear resistor offered in Figure 1.8(a) is

$$v(t) = Ri(t) \tag{1.21}$$

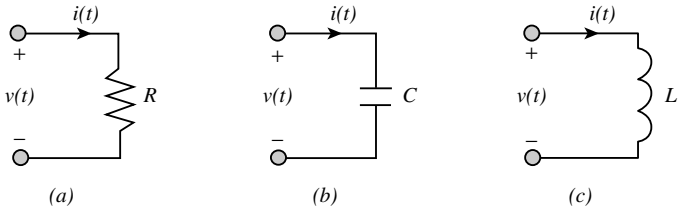


FIGURE 1.8 Circuit schematic symbol and corresponding voltage and current notation for (a) a linear resistor, (b) a linear capacitor, and (c) a linear inductor.

where the voltage, $v(t)$, appearing across the terminals of the resistor and the resultant current, $i(t)$, conducted by the resistor are in associated reference polarity. The resistance, R, is independent of either $v(t)$ or $i(t)$. From (1.14), the dissipated resistor power, which is mainfested in the form of heat, is

$$p_r(t) = i^2(t)R = \frac{v^2(t)}{R} \tag{1.22}$$

The linear capacitor and the linear inductor, with schematic symbols that appear, respectively, in Figures 1.8(b) and (c), store energy as opposed to dissipating power. Their volt-ampere equations are the linear relationships

$$i(t) = C\frac{dv(t)}{dt} \tag{1.23}$$

for the capacitor, whereas for the inductor in Figure 1.8(c),

$$v(t) = L\frac{di(t)}{dt} \tag{1.24}$$

Observe from (1.23) and (1.14) that the power, $p_c(t)$, delivered to the linear capacitor is

$$p_c(t) = v(t)i(t) = Cv(t)\frac{dv(t)}{dt} \tag{1.25}$$

From (1.12), this power is related to the energy, e.g., $w_c(t)$, stored in the form of charge deposited on the plates of the capacitor by

$$Cv(t)dv(t) = dw_c(t) \tag{1.26}$$

It follows that the energy delivered to, and hence stored in, the capacitor from time $t = 0$ to time t is

$$w_c(t) = \frac{1}{2}Cv^2(t) \tag{1.27}$$

It should be noted that this stored energy, like the energy associated with a signal source or a battery voltage, is available to supply power to other elements in the network in which the capacitor is embedded. For example, if very little current is conducted by the capacitor in question, (1.23) implies that the voltage across the capacitor is essentially constant. However, an element whose terminal voltage is a constant and in which energy is stored and therefore available for use behaves as a battery.

If the preceding analysis is repeated for the inductor of Figure 1.8(c), it can be shown that the energy, $w_l(t)$, stored in the inductive element form time $t = 0$ to time t is

$$w_l(t) = \frac{1}{2} L i^2(t) \tag{1.28}$$

Although an energized capacitor conducting almost zero current functions as a voltage source, an energized inductor supporting almost zero terminal voltage emulates a constant current source.

Active vs. Passive

An electrical element or network is said to be *passive* if the power delivered to it, defined in accordance with (1.14), is positive. This definition exploits the requirement that the terminal voltage, $v(t)$, and the element current $i(t)$, appearing in (1.14) be in associated reference polarity. In constrast, an element or network to which the delivered power is negative is said to be *active;* that is, an active element or network generates power instead of dissipating it.

Conventional two-terminal resistors, capacitors, and inductors are **passive elements**. It follows that networks formed of interconnected two-terminal resistors, capacitors, and inductors are **passive networks**. Two-terminal voltage and current sources generally behave as active elements. However, when more than one source of externally applied energy is present in an electrical network, it is possible for one more of these sources to behave as passive elements. Comments similar to those made in conjunction with two-terminal voltage and current sources apply equally well to each of the four possible dependent generators. Accordingly, multiterminal configurations, whose models exploit dependent sources, can behave as either passive or active networks.

Time Varying vs. Time Invariant

The elements of a circuit are defined electrically by an identifying parameter, such as resistance, capacitance, inductance, and the gain factors associated with dependent voltage or current sources. An element whose indentifying parameter changes as a function of time is said to be a **time varying** element. If said parameter is a constant over time, the element in question is **time invariant**. A network containing at least one time varying electrical element it is said to be a time varying network. Otherwise, the network is time invariant. Excluded from the list of elements whose electrical character establishes the time variance or time invariance of a considered network are externally applied voltage and current sources. Thus, for example, a network with internal elements that are exclusively time-invariant resistors, capacitors, inductors, and dependent sources, but which is excited by a sinusoidal signal source, is nonetheless a time-invariant network.

Although some circuits, and particularly electromechanical networks, are purposely designed to exhibit time varying volt–ampere characteristics, **parametric time variance** is generally viewed as a parasitic phenomena in the majority of practical circuits. Unfortunately, a degree of parametric time variance is unavoidable in even those circuits that are specifically designed to achieve input–output response properties that closely approximate time-invariant characteristics. For example, the best of network elements exhibit a slow aging phenomenon that shifts the values of its intrinsic physical parameters. The upshot of these shifts is electrical circuits where overall performance deterioriates with time.

Lumped vs. Distributed

Electrons in conventional conductive elements are not transported instantaneously across elemental cross sections, but their transport velocities are very high. In fact, these velocities approach the speed of light, say c, which is (3×10^8) m/s or about 982 ft/μsec. Electrons and holes in semiconductors are transported at somewhat slower speeds, but generally no less than an order of magnitude or so smaller than the speed of light. The time required to transport charge from one terminal of a two-terminal electrical element to its other terminal, compared with the time required to propagate energy uniformly through the element, determines whether an element is lumped or distributed. In particular, if the time required to transport charge through an element is significantly smaller than the time required to propagate the

energy through the element that is required to incur such charge transport, the element in question is said to be lumped. On the other hand, if the charge transport time is comparable to the energy propagation time, the element is said to be **distributed**.

The concept of a lumped, as opposed to a distributed, circuit element can be qualitatively understood through a reconsideration of the circuit provided in Figure 1.7. As argued, the indicated element current, $i(t)$, is identical to the indicated source current, $i_s(t)$. This equality implies that $i(t)$, is effectively circulating around the loop that is electrically formed by the interconnection of the signal source voltage, $v_s(t)$, to the element. Equivalently, the subject equality implies that $i(t)$ is entering terminal 1 of the element and simultaneously is exiting at terminal 2, as illustrated. Assuming that the element at hand is not a semiconductor, the current, $i(t)$, arises from the transport of electrons through the element in a direction opposite to that of the indicated polarity of $i(t)$. Specifically, electrons must be transported from terminal 2, at the bottom of the element, to terminal 1, at the top of the element, and in turn the requisite amount of energy must be applied in the immediate neighborhoods of both terminals. The implication of presuming that the element at hand is lumped is that $i(t)$ is entering terminal 1 at precisely the same time that it is leaving terminal 2. Such a situation is clearly impossible, for it mandates that electrons be transported through the entire length of the element in zero time. However, given that electrons are transported at a nominal velocity of 982 ft/μsec, a very small physical elemental length renders the approximation of zero electron transport time reasonable. For example, if the element is 1/2 inch long (a typical size for an off-the-shelf resistor), the average transport time for electrons in this unit is only about 42.4 psec. As long as the period of the applied excitation, $v_s(t)$, is significantly larger than 42.4 psec, the electron transport time is significantly smaller than the time commensurate with the propagation of this energy through the entire element. A period of 42.4 psec corresponds to a signal whose frequency of approximately 23.6 GHz. Thus, a 1/2-in resistive element excited by a signal whose frequency is significantly smaller than 23.6 GHz can be viewed as a lumped circuit element.

In the vast majority of electrical and electronic networks it is difficult not to satisfy the lumped circuit approximation. Nevertheless, several practical electrical systems cannot be viewed as lumped entities. For example, consider the lead-in wire that connects the antenna input terminals of a frequency modulated (FM) radio receiver to an antenna, as diagrammed in Figure 1.9. Let the signal voltage, $v_a(t)$, across the lead-in wires at point "a" be the sinusoid,

$$v_a(t) = V_M \cos(\omega t) \tag{1.29}$$

where V_M represent the amplitude of the signal, and ω is its frequency in units of radians per second. Consider the case in which $\omega = 2\pi(103.5 \times 10^6)$ rad/s, which is a carrier frequency lying within the commercial FM broadcast band. This high signal frequency makes the length of antenna lead-in wire critically important for proper signal reception.

In an attempt to verify the preceding contention, let the voltage developed across the lead-in lines at point "b" in Figure (1.9) be denoted as $v_b(t)$, and let point "b" be 1 foot displaced from point "a"; that is, $L_{ab} = 1$ foot. The time, τ_{ab} required to transport electrons over the indicated length, L_{ab}, is

$$\tau_{ab} = \frac{L_{ab}}{c} = 1.018 \text{ ns} \tag{1.30}$$

Thus, assuming an idealized line in the sense of zero effective resistance, capacitance, and inductance, the signal, $v_b(t)$, at point "b" is seen as the signal appearing at "a", delayed by approximately 1.02 ns. It follows that

$$v_b(t) = V_M \cos\left[\omega\left(t - \tau_{ab}\right)\right] = V_M \cos(\omega t - 0.662) \tag{1.31}$$

where the phase angle associated with $v_b(t)$ is 0.662 radian, or almost 38°. Obviously, the signal established at point "b" is a significantly phase-shifted version of the signal presumed at point "a".

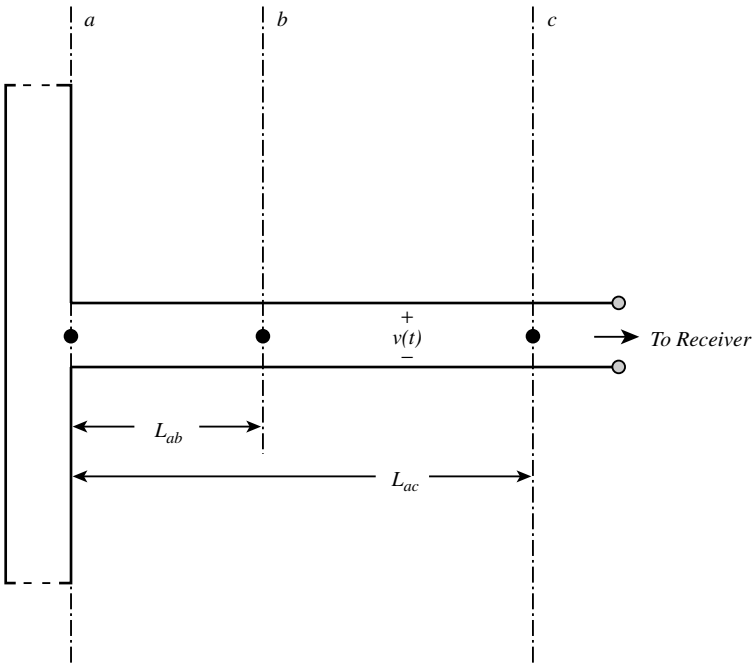


FIGURE 1.9 Schematic abstraction of a dipole antenna for an FM receiver application.

An FM receiver can effectively retrieve the signal voltage, $v_a(t)$, by detecting a phase-inverted version of $v_a(t)$ at its input terminals. To this end, it is of interest to determine the length, L_{ac}, such that the signal, $v_c(t)$, established at point "c" in Figure 1.9 is

$$v_c(t) = V_M \cos(\omega t - \pi) \tag{1.32}$$

The required phase shift of 180°, or π radians, corresponds to a time delay, τ_{ac}, of

$$\tau_{ac} = \frac{\pi}{\omega} = 4.831 \text{ ns} \tag{1.33}$$

In turn, a time delay of τ_{ac} implies a required line length, L_{ac} of

$$L_{ac} = c\tau_{ac} = 4.744 \text{ ft} \tag{1.34}$$

A parenthetically important point is the observation that the carrier frequency of 103.5 MHz corresponds to a wavelength, λ, of

$$\lambda = \frac{2\pi c}{\omega} = 9.489 \text{ ft} \tag{1.35}$$

Accordingly, the lead-in length computed in (1.34) is $\lambda/2$; that is, a half-wavelength.

2

Network Laws and Theorems

Ray R. Chen
San Jose State University

Artice M. Davis
San Jose State University

Marwan A. Simaan
University of Pittsburgh

2.1 Kirchhoff's Voltage and Current Laws

Ray R. Chen and Artice M. Davis

Circuit analysis, like Euclidean geometry, can be treated as a mathematical system; that is, the entire theory can be constructed upon a foundation consisting of a few fundamental concepts and several axioms relating these concepts. As it happens, important advantages accrue from this approach — it is not simply a desire for mathematical rigor, but a pragmatic need for simplification that prompts us to adopt such a mathematical attitude.

The basic concepts are conductor, element, time, voltage, and current. Conductor and element are axiomatic; thus, they cannot be defined, only explained. A conductor is the idealization of a piece of copper wire; an element is a region of space penetrated by two conductors of finite length termed **leads** and pronounced "leeds". The ends of these leads are called **terminals** and are often drawn with small circles as in Figure 2.1.

Conductors and elements are the basic objects of circuit theory; we will take time, voltage, and current as the basic variables. The time variable is measured with a clock (or, in more picturesque language, a chronometer). Its unit is the second, *s*. Thus, we will say that time, like voltage and current, is defined **operationally**, that is, by means of a measuring instrument and a procedure for measurement. Our view of reality in this context is consonant with that branch of philosophy termed **operationalism** [1].

Voltage is measured with an instrument called a *voltmeter,* as illustrated in Figure 2.2. In Figure 2.2, a voltmeter consists of a readout device and two long, flexible conductors terminated in points called **probes** that can be held against other conductors, thereby making electrical contact with them. These conductors are usually covered with an insulating material. One is often colored red and the other black. The one colored red defines the positive polarity of voltage, and the other the negative polarity. Thus, voltage is always measured between two conductors. If these two conductors are element leads, the voltage is that across the corresponding element. Figure 2.3 is the symbolic description of such a measurement; the variable v, along with the corresponding plus and minus signs, means exactly the experimental procedure depicted in Figure 2.2, neither more nor less. The outcome of the measurement, incidentally, can be either positive or negative. Thus, a reading of $v = -12$ V, for example, has meaning only when

FIGURE 2.1 Conductors and elements.

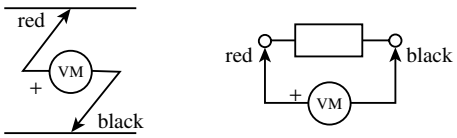


FIGURE 2.2 The operational definition of voltage.

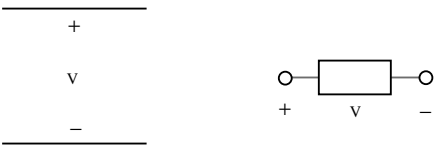


FIGURE 2.3 The symbolic description of the voltage measurement.

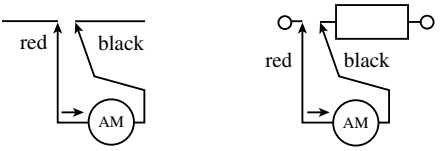


FIGURE 2.4 The operational definition of current.

FIGURE 2.5 The symbolic representation of a current measurement.

viewed within the context of the measurement. If the meter leads are simply reversed after the measurement just described, a reading of $v' = +12$ V will result. The latter, however, is a different variable; hence, we have changed the symbol to v'. The V after the numerical value is the unit of voltage, the volt, V.

Although voltage is measured across an element (or between conductors), current is measured through a conductor or element. Figure 2.4 provides an operational definition of current. One cuts the conductor or element lead and touches one meter lead against one terminal thus formed and the other against the second. A shorthand symbol for the meter connection is an arrow close to one lead of the ammeter. This arrow, along with the meter reading, defines the current. We show the shorthand symbol for a current in Figure 2.5. The reference arrow and the symbol i are shorthand for the complete measurement in Figure 2.4 — merely this and nothing more. The variable i can be either positive or negative; for example, one possible outcome of the measurement might be $i = -5$ A. The A signifies the unit of current, the ampere. If the red and black leads in Figure 2.4 were reversed, the reading sign would change.

Table 2.1 provides a summary of the basic concepts of circuit theory: the two basic objects and the three fundamental variables. Notice that we are a bit at variance with the SI system here because although time and current are considered fundamental in that system, voltage is not. Our approach simplifies things, however, for one does not require any of the other SI units or dimensions. All other quantities

TABLE 2.1 Summary of the Basic Concepts of Circuit Theory

Objects		Variables		
Conductor	Element	Time	Voltage	Current
—	—	Seconds, s	Volt, V	Ampere, A

are derived. For instance, charge is the integral of current and its unit is the ampere-second, or the coulomb, C. Power is the product of voltage and current. Its unit is the watt, W. Energy is the integral power, and has the unit of the watt-second, or joule, J. In this manner one avoids the necessity of introducing mechanical concepts, such as mechanical work, as being the product of force and distance.

In the applications of circuit theory, of course, one has need of the other concepts of physics. If one is to use circuit analysis to determine the efficiency of an electric motor, for example, the concept of mechanical work is necessary. However — and this is the main point of our approach — the introduction of such concepts is not essential in the analysis of a circuit itself. This idea is tied in to the concept of modeling. The basic catalog of elements used here does not include such things as temperature effects or radiation of electromagnetic energy. Furthermore, a "real" element such as resistor is not "pure." A real resistor is more accurately modeled, for many purposes, as a resistor plus series inductance and shunt capacitance. The point is this: In order to adequately model the "real world" one must often use complicated combinations of the basic elements. Additionally, to incorporate the influence of variables such as temperature, one must assume that certain parameters (such as resistance or capacitance) are functions of that variable. It is the determination of the more complicated model or the functional relationship of a given parameter to, for example, temperatures that fall within the realm of the practitioner. Such ideas were discussed more fully in Chapter 1. Circuit analysis merely provides the tools for analyzing the end result.

The radiation of electromagnetic energy is, on the other hand, a quite different aspect of circuit theory. As will be seen, circuit analysis falls within a regime in which such behavior can be neglected. Thus, the theory of circuit analysis we will expound has a limited range of application: to low frequencies or, what is the same in the light of Fourier analysis, to waveforms that do not vary too rapidly.

We are now in a position to state two basic axioms, which we will assume all circuits obey:

Axiom 1: The behavior of an element is completely determined by its *v–i* characteristic, which can be determined by tests made on the element in isolation from the other elements in the circuit in which it is connected.

Axiom 2: The behavior of a circuit is independent of the size or the shape or the orientation of its elements, the conductors that interconnect them, and the element leads.

At this point, we loosely consider a circuit to be any collection of elements and conductors, although we will sharpen our definition a bit later. Axiom 1 means that we can run tests on an element in the laboratory, then wire it into a circuit and have the assurance that it will not exhibit any new and different behavior. Axiom 2 means that it is only the *topology* of a circuit that matters, not the way the circuit is stretched or bent or rearranged, so long as we do not change the listing of which element leads are connected to which others or to which conductors.

The remaining two axioms are somewhat more involved and require some discussion of circuit topology. Consider, for a moment, the collection of elements in Figure 2.6. We labeled each element with a letter to distinguish it from the others. First, notice the two solid dots. We refer to them as **joints**. The idea is that they represent "solder joints," where the ends of two or more leads or conductors were connected. If only two ends are connected we do not show the joints explicitly; where three or more are connected, however, they are drawn. We temporarily (as a test) erase all of the element bodies and replace them with open space. The result is given in Figure 2.7. We refer to each of the interconnected "islands" of a conductor as a **node**. This example circuit has six nodes, and we labeled them with the numbers one through six for identification purposes.

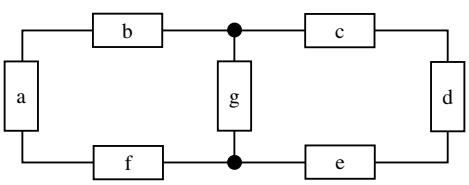


FIGURE 2.6 An example circuit.

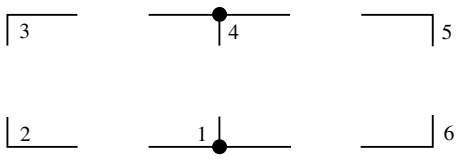


FIGURE 2.7 The nodes of the example circuit.

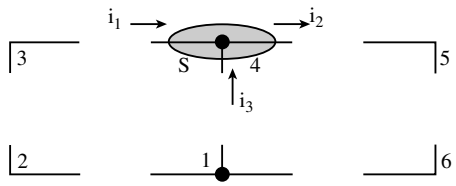


FIGURE 2.8 Illustration of Kirchhoff's current law.

Axiom 3 (Kirchhoff's Current Law): The charge on a node or in an element is identically zero at all instants of time.

Kirchhoff's current law (KCL) is not usually phrased in quite this manner. Thus, let us consider the closed (or "Gaussian") surface S in Figure 2.8. We assume that it is penetrated only by conductors. The elements, of course, are there; we simply do not show them so that we can concentrate on the conductors. We have arbitrarily defined the currents in the conductors penetrating *S*. Now, recalling that charge is the time integral of the current and thus has the same direction as the current from which it is derived, one can phrase Axiom 3 as follows:

$$\sum_s q_{\text{in}} = q_{\text{enclosed}} = 0 \tag{2.1}$$

at each instant of time. This equation is simply one form of conservation of charge. Because current is the time derivative of voltage, one can also state that

$$\sum_s i_{\text{in}} = 0 \tag{2.2}$$

at each and every time instant. This last equation is the usual phrasing of KCL. The subscript "in" means that a current reference pointed inward is to be considered positive; by default, therefore, a current with its reference pointed outward is to have a negative sign affixed. This sign is in addition to any negative sign that might be present in the value of each variable. For node 4 in Figure 2.8, KCL in its current form, therefore, reads

$$i_1 - i_2 + i_3 = 0 \tag{2.3}$$

Two other ways of expressing KCL (in current form) are

$$\sum_s i_{\text{out}} = 0 \tag{2.4}$$

and

$$\sum_s i_{\text{in}} = \sum_s i_{\text{out}} \tag{2.5}$$

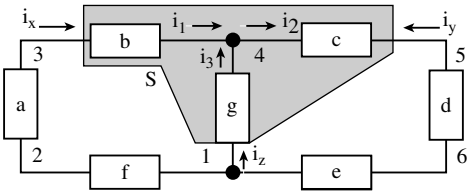


FIGURE 2.9 KCL for a more general surface.

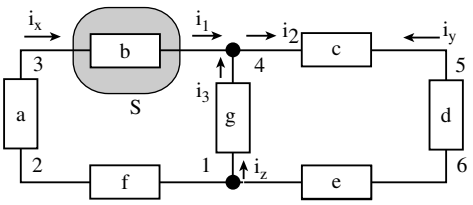


FIGURE 2.10 KCL for a single element.

The equivalent charge forms are also clearly valid. We emphasize the latter to a greater extent than is usual in the classical treatment because of the current interest in charge distribution and transfer circuits.

The Gaussian surface used to express KCL is not constrained to enclose only conductors. It can enclose elements as well, although it still can be penetrated by only conductors (which can be element leads). Thus, consider Figure 2.9, which illustrates the same circuit with which we have been working. Now, however, the elements are given and the Gaussian surface encloses three elements, as well as conductors carrying the currents previously defined. Because these currents are not carried in the conductors penetrating the surface under consideration, they do not enter into KCL for that surface. Instead, KCL becomes

$$i_x + i_y + i_z = 0 \tag{2.6}$$

As a special case let us look once more at the preceding figure, but use a different surface, one enclosing only the element *b*. This is depicted in Figure 2.10. If we refer to Axiom 3, which notes that charge cannot accumulate inside an element, and apply charge conservation, we find that

$$i_x = i_1 \tag{2.7}$$

This states that the current into any element in one of its leads is the same as the current leaving in the other lead. In addition, we see that KCL for nodes and KCL for elements (both of which are implied by Axiom 3) imply that KCL holds for any general closed surface penetrated only by conductors such as the one used in connection with Figure 2.9.

In order to phrase our last axiom, we must discuss circuit topology a bit more, and we will continue to use the circuit just considered previously. We define a **path** to be an ordered sequence of elements having the property that any two consecutive elements in the sequence share a common node. Thus, referring for convenience back to Figure 2.10, we see that {*f, a, b*} is a path. The elements *f* and *a* share node 2 and *a* and *b* share node 3. One lead of the last element in a path is connected to a node that is not shared with the preceding element. Such a node is called the **terminal** node of the path. Similarly, one lead of the first element in the sequence is connected to a node that is not shared with the preceding element.[1] It is called the *initial* node of the path. Thus, in the example just cited, node 1 is the initial node and node 4 is the final node. Thus, a direction is associated with a path, and we can indicate it diagram-

[1]We assume that no element has its two leads connected together and that more than two elements are in the path in this definition.

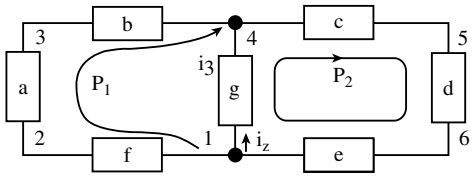


FIGURE 2.11 Circuit paths.

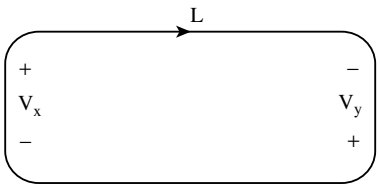


FIGURE 2.12 Voltage rise and drop.

matically by means of an arrow on the circuit. This is illustrated in Figure 2.11 for the path $P_1 = \{f, a, b\}$ and $P_2 = \{g, c, d, e\}$.

If the initial node is identical to the terminal node, then the corresponding path is called a *loop*. An example is $\{f, a, b, g\}$. The patch P_2 is a loop. An alternate definition of a loop is as a collection of branches having the property that each node connected to a patch branch is connected to precisely two path branches; that is, it has **degree two** relative to the path branches.

We can define the voltage across each element in our circuit in exactly two ways, corresponding to the choices of which lead is designated plus and which is designated minus. Figure 2.12 presents two voltages and a loop L in a highly stylized manner. We have purposely not drawn the circuit itself so that we can concentrate on the essentials in our discussion. If the path enters the given element on the lead carrying the minus and exits on the one carrying the positive, its voltage will be called a **voltage rise**; however, if it enters on the positive and exits on the minus, the voltage will be called a **voltage drop**. If the signs of a voltage are reversed and a negative sign is affixed to the voltage variable, the value of that variable remains unchanged; thus, note that a negative rise is a drop, and vice versa.

We are now in a position to state our fourth and final axiom:

Axiom 4 (Kirchhoff's Voltage Law): The sum of the voltage rises around any loop is identically zero at all instants of time.

We refer to this law as KVL for the sake of economy of space. Just as KCL was phrased in terms of charge, KVL could just as well be phrased in terms of **flux linkage**. Flux linkage is the time integral of voltage, so it can be said that the sum of the flux linkages around a loop is zero. In voltage form, we write

$$\sum_{\text{loop}} v_{\text{rises}} = 0 \tag{2.8}$$

We observed that a negative rise is a drop, so

$$\sum_{\text{loop}} v_{\text{drops}} = 0 \tag{2.9}$$

or

$$\sum_{\text{loop}} v_{\text{rises}} = \sum_{\text{loop}} v_{\text{drops}} \tag{2.10}$$

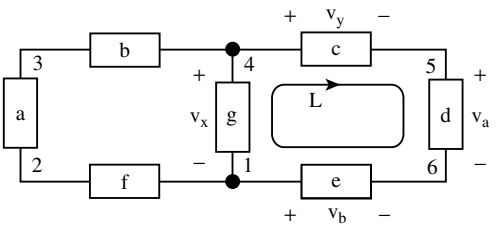


FIGURE 2.13 Illustration of Kirchhoff's voltage law.

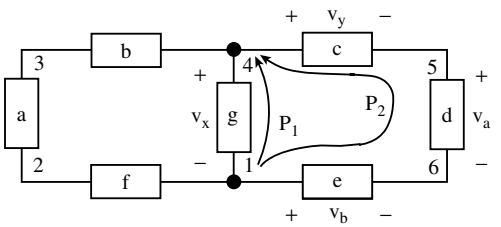


FIGURE 2.14 Path form of KVL.

Thus, in Figure 2.13, we could write [should we choose to use the form of (2.8)]

$$v_x - v_y - v_a + v_b = 0 \tag{2.11}$$

Clearly, one can rearrange KVL into many different algebraic forms that are equivalent to those just stated; one form, however, is more useful in circuit computations than many others. It is known as the **path form** of KVL. To better appreciate this form, review Figure 2.13. This time, however, the paths are defined a bit differently. As illustrated in Figure 2.14, we consider two paths, P_1 and P_2, having the same initial and terminal nodes, 1 and 4, respectively.[2] We can rearrange (2.11) into the form

$$v_x = -v_b + v_a + v_y \tag{2.12}$$

This form is often useful for finding one unknown voltage in terms of known voltages along some given path. In general, if P_1 and P_2 are two paths having the same initial and final nodes,

$$\sum_{P_1} v_{\text{rises}} = \sum_{P_2} v_{\text{rises}} = 0 \tag{2.13}$$

Be careful to distinguish this equation from (2.10). In the present case two paths are involved; in the former we find only a single loop, and drops are located on one side of the equation and rises on the other. One might call the path form the "all roads lead to Rome" form.

We covered four basic axioms, and these are all that are needed to construct a mathematical theory of circuit analysis. The first axiom is often referred to by means of the phrase "lumped circuit analysis", for we assume that all the physics of a given element are internal to that element and are of no concern to us; we are only interested in the *v–i* characteristic. That is, we are treating all the elements as lumps of matter that interact with the other elements in a circuit by means of the voltage and current at their leads. The second axiom says that the physical construction is irrelevant and that the interconnections are completely described by means of the circuit graph. Kirchhoff's current law is an expression of conservation of charge, plus the assumption that neither conductors nor elements can maintain a net

[2]If one defines the **negative** of a path as a listing of the same elements as the original path in the reverse order and summation of two paths as a concatenation of the two listings, one sees that $P_1 - P_2 = L$, the loop in Figure 2.13.

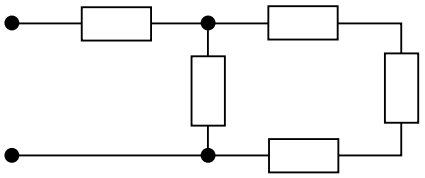


FIGURE 2.15 A "noncircuit."

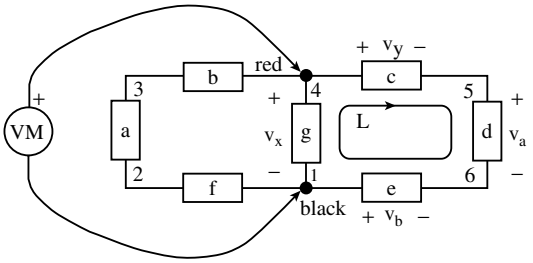


FIGURE 2.16 Node voltages.

charge. In this connection, observe that a capacitor maintains a charge separation internally, but it is a separation of two charges of opposite sign; thus, the total algebraic charge within it is zero. Finally, KVL is an expression of conservation of flux linkage. If $\mathrm{l}(t) = \int_{-\infty}^{t} v(\alpha)d\alpha$ is the flux linkage, then one can write[3] (using one form of KVL)

$$\sum_{\text{loop}} \lambda_{\text{rises}}^{(t)} = 0 \tag{2.14}$$

In the theory of electromagnetics, one finds that this equation does not hold exactly; in fact, the right-hand side is equal to the negative of the derivative of the magnetic flux contained within the loop (this is the Faraday–Lenz law). If, however, the time variation of all signals in the circuit are slow, then the right-hand side is approximately zero and KVL can be assumed to hold. A similar result holds also for KCL. For extremely short instants of time, a conductor can support an unbalanced charge. One finds, however, that the "relaxation" time of such unbalanced charge is quite short in comparison with the time variations of interest in the circuits considered in this text.

Finally, we tie up a loose end left hanging at the beginning of this subsection. We consider a circuit to be, not just any collection of elements that are interconnected, but a collection having the property that each element is contained in at least one loop. Thus, the circuit in Figure 2.15 is not a circuit; instead, it must be treated as a *subcircuit*, that is, as part of a larger circuit in which it is to be imbedded.

The remainder of this section develops the application of the axioms presented here to the analysis of circuits. The reader is referred to [2, 3, 4] for a more detailed treatment.

Nodal Analysis

Nodal analysis of electric circuits, although using all four of the fundamental axioms presented in the introduction, concentrates upon KCL explicitly. Kirchhoff's voltage law is also satisfied automatically in view of the way the basic equations are formulated. This effective method uses the concept of a **node voltage**. Figure 2.16 illustrates the concept. Observe a voltmeter, with its black probe attached to a single node, called the **reference node**, which remains fixed during the course of the investigation. In the case

[3]One might anticipate a constant on the right side of (2.14); however, a closer investigation reveals that it is more realistic and pragmatic to assume that all signals are one-sided and that all elements are causal. This implies that the constant is zero. Two-sided signals only arise legitimately within the context of steady-state behavior of stable circuits and systems.

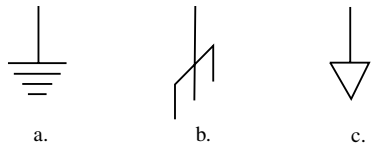


FIGURE 2.17 Reference node symbols.

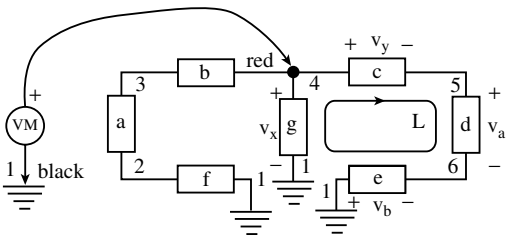


FIGURE 2.18 An alternate drawing.

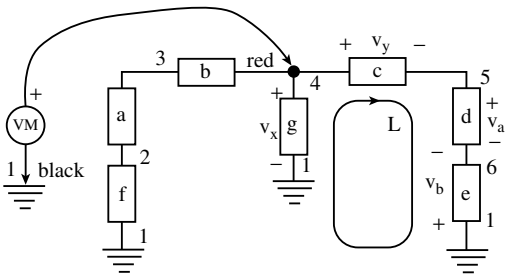


FIGURE 2.19 An equivalent drawing.

shown node 1 is the reference node. The red probe is shown being touched to node 4; therefore, we call the resulting voltage v_4. The subscript denotes the node and the result is always assumed to have its positive reference on the given node. In the present instance v_4 is identical to the element voltage because element g (across which v_g is defined) is connected between node 4 and the reference node. Note that the voltage of such an element is always either the node voltage or its negative, depending upon the reference polarities of its associated element voltage. If we were to touch the red probe to node 5, however, no element voltage would have this relationship to the resulting node voltage v_5 because no elements are connected directly between nodes 5 and 1.

The concept of reference node is used so often that a special symbol is used for it [see Figure 2.17(a)]; alternate symbols often seen on circuit diagrams are shown in the figure as well. Often one hears the terms "ground" or "ground reference" used. This is commonly accepted argot for the reference node; however, one should be aware that a safety issue is involved in the *process* of grounding a circuit or appliance. In such cases, sometimes one symbol specifically means "earth ground" and one or more other symbols are used for such things as "signal ground" or "floating ground", although the last term is something of an oxymoron. Here, we use the terms "reference" or "reference node." The reference symbol is quite often used to simplify the drawing of a circuit. The circuit in Figure 2.16, for instance, can be redrawn as in Figure 2.18; circuit operation will be unaffected. Note that all four of the reference symbols refer to a single node, node 1, although they are shown separated from one another. In fact, the circuit is not changed electrically if one bends the elements around and thereby separates the ground symbols even more, as we have done in Figure 2.19. Notice that the loop L shown in the original figure, Figure 2.16, remains a loop, as in Figures 2.18 and 2.19. Redrawing a circuit using ground reference symbols does not alter the circuit topology, the circuit graph.

Suppose the red probe were moved to node 5. As described previously, no element is directly connected between nodes 5 and 1; hence, node voltage v_5 is not an element voltage. However, the element voltages

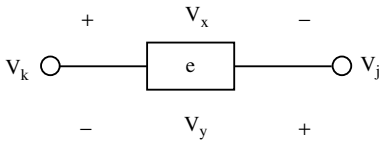


FIGURE 2.20 A floating element and its voltage.

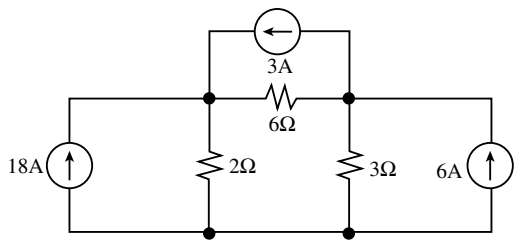


FIGURE 2.21 An example circuit.

and the node voltages are directly related in a one-to-one fashion. To see how, look at Figure 2.20. This figure shows a "floating element," *e*, which is connected between two nodes, *k* and *j*, neither of which is the reference node. It is vital here to remember that all node voltages are assumed to have their positive reference polarities on the nodes themselves and their negative reference on the reference node. Now, we can define the element voltage in either of two possible ways, as illustrated in the figure. Kirchhoff's voltage law (the simplest form perhaps being the path form) shows at once that

$$v_x = v_k - v_j \tag{2.15}$$

and

$$v_y = v_j - v_k \tag{2.16}$$

An easy mnemonic for this result is the following:

$$v_{\text{floatingelement}} = v_+ - v_- \tag{2.17}$$

where v_+ is the node voltage of the node to which the element lead associated with the positive reference for the element voltage is connected, and v_- is the node voltage of the node to which the lead carrying the negative reference for the element voltage is connected. We refer to an element that is not floating, by the way, as being "grounded."

It is easy to see that a circuit having N nodes has $N-1$ node voltages; further, if one uses (2.17), any element voltage can be expressed in terms of these $N-1$ node voltages. Then, for any invertible element,[4] one can determine the element current. The nodal analysis method uses this fact and considers the node voltages to be the unknown variables.

To illustrate the method, first consider a resistive circuit that contains only resistors and/or independent sources. Furthermore, we initially limit our investigation to circuits whose only independent sources (if any) are current sources. Such a circuit is depicted in Figure 2.21. Because nodal analysis relies upon the node voltages as unknowns, one must first select an arbitrary node for the reference. For circuits that contain voltage sources, one can achieve some simplification for hand analysis by choosing the reference wisely; however, if current sources are the only type of independent source present, one can choose it arbitrarily. As it happens, physical intuition is almost always better served if one chooses the bottom

[4]For instance, resistors, capacitors, and inductors are invertible in the sense that one can determine their element currents if their element voltages are known.

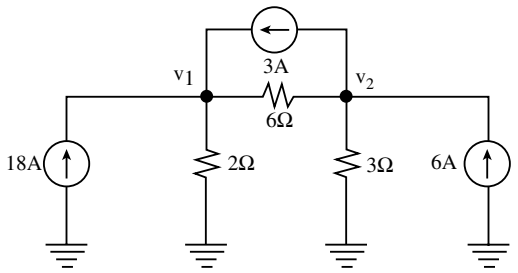


FIGURE 2.22 The example circuit prepared for nodal analysis.

node. Such is done here and the circuit is redrawn using reference symbols as in Figure 2.22. Here, we have arbitrarily assigned node voltages to the $N - 1 = 2$ nonreference nodes. In performing these two steps, we have "prepared the circuit for nodal analysis." The next step involves writing one KCL equation at each of the nonreference nodes. As it happens, the resulting equations are nice and compact if the form

$$\sum_{\text{node}} i_{\text{out}}\left(R's\right) = \sum_{\text{node}} i_{\text{in}}\left(I - \text{sources}\right) \tag{2.18}$$

is used. Here, we mean that the currents leaving a node through the resistors must sum up to be equal to the current being supplied to that node from current sources. Because these two types of elements are exhaustive for the circuits we are considering, this form is exactly equivalent to the other forms presented in the introduction. Furthermore, for a current leaving a node through a resistor, the floating element KVL result in (2.17) is used along with Ohm's law:

$$\sum_{j=1}^{N-1} \frac{v_k - v_j}{R_{kj}} = \sum_{q=1}^{M_k} i_{sq}\left(\text{node } k\right). \tag{2.19}$$

In this equation for node k, R_{kj} is the resistance between nodes k and j (or the equivalent resistance of the parallel combination if more than one are found), i_{sq} is the value of the qth current source connected to node k (positive if its reference is toward node k), and M_k is the number of such sources. Clearly, one can simply omit the $j = k$ term on the left side because $v_k - v_k = 0$.

The nodal equations for our example circuit are

$$\frac{v_1}{2} + \frac{v_1 - v_2}{6} = 18 + 3 \tag{2.20}$$

and

$$\frac{v_2 - v_1}{6} + \frac{v_2}{3} = 6 - 3 \tag{2.21}$$

Notice, by the way, that we are using units of A, Ω, and *V*. It is a simple matter to show that KVL, KCL, and Ohm's law remain invariant if we use the consistent units of mA, kΩ, and V. The latter is often a more practical system of units for filter design work. In the present case the matrix form of these equations is

$$\begin{bmatrix} \frac{2}{3} & -\frac{1}{6} \\ -\frac{1}{6} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} 21 \\ 3 \end{bmatrix} \tag{2.22}$$

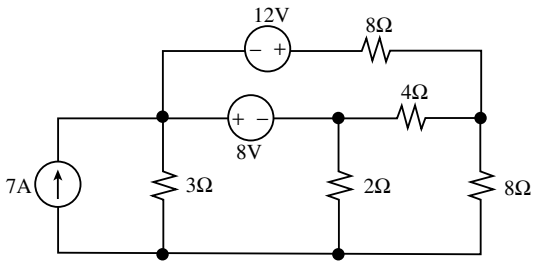


FIGURE 2.23 An example circuit.

It can be verified easily that the solution is $v_1 = 36$ V and $v_2 = 18$ V. To see that one can compute the value of any desired variable from these two voltages, consider the problem of determining the current i_6 (let us call it) through the horizontal 6 Ω resistor from right to left. One can simply use the equation

$$i_6 = \frac{v_2 - v_1}{6} = \frac{18 - 36}{6} = -3\text{A} \tag{2.23}$$

The previous procedure works for essentially all circuits encountered in practice. If the coefficient matrix on the left in (2.22) (which will always be symmetric for circuits of the type we are considering) is nonsingular, a solution is always possible. It is surprisingly difficult, however, to determine conditions on the circuit under which solutions do not exist, although this is discussed at greater length in a later subsection.

Suppose, now, that our circuit to be solved contains one or more independent voltage sources in addition to resistors and/or current sources. This constrains the node voltages because a given voltage source value must be equal to the difference between two node voltages if it is floating and to a node voltage or its negative if it is grounded. One might expect that this complicates matters, but fortunately the converse is true.

To explore this more fully, examine the example circuit in Figure 2.23. The algorithm just presented will not work as is because it relies upon balancing the current between resistors and current sources. Thus, it seems that we must account in some fashion for the currents in the voltage sources. In fact, we do not, as the following analysis shows. The key step in our reasoning is this: the analysis *procedure* should not depend upon the values of the independent circuit variables, that is, on the values of the currents in the current sources and voltages across the voltage sources. This is almost inherent in the definition of an independent source, for it can be adjusted to any value whatsoever. What we are assuming in addition to this is simply that we would not write one given set of equations for a specific set of source values, then change to another set of equations when these values are altered. Thus, let us test the circuit by temporarily deactivating all the independent sources (i.e., by making their values zero). Recalling that a deactivated voltage source is equivalent to a short circuit and a deactivated current source to an open circuit, we have the resulting configuration of Figure 2.24. The resulting nodes are shaded for convenience. Note carefully, however, that the nodes in the circuit under test are not the same as those in the original

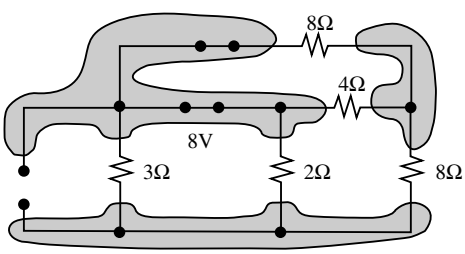


FIGURE 2.24 The example circuit deactivated.

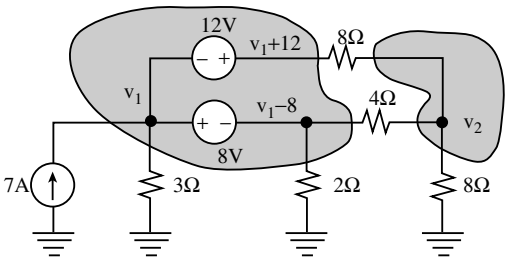


FIGURE 2.25 The example circuit prepared for nodal analysis.

circuit, although they are related. Notice that, for the circuit under test, all the resistor voltages would be determined by the node voltages as expected; however, *the number of nodes has been reduced by one for each voltage source*. Hence, we suspect that the required number of KCL equations N_{ne} (and the number of independent node voltages) is

$$N_{ne} = N - 1 - N_v \tag{2.24}$$

where N_v is the number of voltage sources. In the example circuit one can easily compute this required number to be 5 –1 –2 = 2. This is compatible with the fact that clearly three nodes (3 – 1 = 2 nonreference nodes) are clearly found in Figure 2.24.

It should also be rather clear that there is only one independent voltage within each of the shaded regions shown in Figure 2.24. We can use KVL to express any other in terms of that one. For example, in Figure 2.25 we have redrawn our example circuit with the bottom node arbitrarily chosen as the reference. We have also arbitrarily chosen a node voltage within the top left surface as the unknown v_1. Note how we have used KVL (the path form, again, is perhaps the most effective) to determine the node voltages of all the other nodes within the top left surface. Any set of connected conductors, leads, and voltage sources to which only one independent voltage can be assigned is called a **generalized node**. If that generalized node does not include the reference node, it is termed a **supernode**. The node within the shaded surface at the top left in Figure 2.25, however, has no voltage sources; hence, it is called an **essential node**.

As pointed out earlier, the equations that one writes should not depend upon the values of the independent sources. If one were to reduce all the independent sources to zero, each generalized node would reduce to a single node; hence, only one equation should be written for each supernode. One equation should be written for essential node also; it is unaffected by deactivation of the independent sources. Observe that deactivation of the current sources does not reduce the number of nodes in a circuit.

Writing one KCL equation for the supernode and one for the essential node in Figure 2.25 results in

$$\frac{v_1}{3} + \frac{v_1 - 8}{2} + \frac{v_1 - 8 - v_2}{4} + \frac{v_1 + 12 - v_2}{8} = 7 \tag{2.25}$$

and

$$\frac{v_2}{8} + \frac{v_2 - (v_1 - 8)}{4} + \frac{v_2 - (v_1 + 12)}{8} = 0 \tag{2.26}$$

In matrix form, one has

$$\begin{bmatrix} \frac{29}{24} & -\frac{3}{8} \\ -\frac{3}{8} & \frac{1}{2} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \end{bmatrix} = \begin{bmatrix} \frac{23}{2} \\ -\frac{1}{2} \end{bmatrix} \tag{2.27}$$

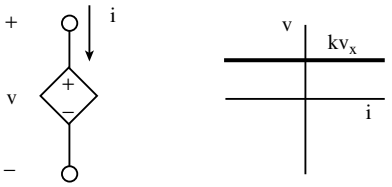


FIGURE 2.26 A dependent source.

The solution is $v_1 = 12$ V and $v_2 = 8$. Notice once again that the coefficient matrix on the left-hand side is symmetric. This actually follows from our earlier observation about this property for circuits containing only current sources and resistors because the voltage sources only introduce knowns into the nodal equations, thus modifying the right-hand side of (2.27).

The general form for nodal equations in any circuit containing only independent sources and resistors, based upon our foregoing development, is

$$A\bar{v}_n = F_v\bar{v}_s + F_I\bar{i}_s \tag{2.28}$$

where A is a symmetric square matrix of constant coefficients, F_v and F_I are rectangular matrices of constants, and $\bar{v}_n$ is the column matrix of independent mode voltages. The vectors $\bar{v}_s$ and $\bar{i}_s$ are column matrices of independent source values. Clearly, if A is a nonsingular matrix, (2.28) can be solved for the node voltages. Then, using KVL and/or Ohm's law, one can solve for any element current or voltage desired. Equally clearly, if a solution exists, it is a multilinear function of the independent source values.[5]

Now suppose that the circuit under consideration contains one or more dependent sources. Recall that the two-terminal characteristics of such elements are indistinguishable from those of the corresponding independent sources except for the fact that their value depends upon some other circuit variable. For instance, in Figure 2.26 a voltage-controlled voltage source (VCVS) is shown. Its v–i characteristic is identical to that of an independent source except for the fact that its voltage (the controlled variable) is a constant multiple[6] of another circuit variable (the controlling variable), in this case another voltage. This fact will be relied upon to develop a modification to the nodal analysis procedure.

We will adopt the following attitude: We will imagine that the dependent relationship, kv_x in Figure 2.26, is a label pasted to the surface of the source in much the same way that a battery is labeled with its voltage. We will imagine ourselves to take a small piece of opaque masking tape and apply it over this label; we will call this process **taping the dependent source**. This means that we are — temporarily — treating it as an independent source. The usual nodal analysis procedure is then followed, which results in (2.28). Then, we imagine ourselves to remove the tape from the dependent source(s) and note that the relationship is linear, with the controlling variables as the independent ones and the controlled variables the dependent ones. We next express the controlling variables — and thereby the controlled ones as well — in terms of the node voltages using KVL, KCL, and Ohm's law. The resulting relationships have the forms

$$\bar{v}_c = B'\bar{v}_n + C'\bar{y}_{si} + D'\bar{i}_{si} \tag{2.29}$$

and

$$\bar{i}_c = B''\bar{v}_n + C''\bar{v}_{si} + D''\bar{i}_{si} \tag{2.30}$$

Here, the subscript i refers to the fact that the corresponding sources are the independent ones. Noting that $\bar{v}_c$ and $\bar{i}_c$ appear on the right-hand side of (2.28) because they are source values, one can use the last two results to express the vectors of all source voltages and all source currents in that equation in the form

[5]That is, it is a linear function of the vector consisting of all of the independent source values.

[6]Thus, one should actually refer to such a device as a *linear* dependent source.

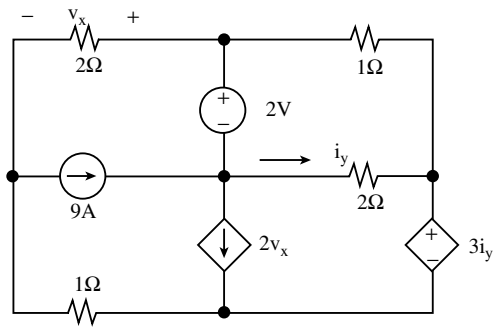


FIGURE 2.27 An example circuit.

$$\bar{v}_c = B'''\bar{v}_{si} + C'''\bar{i}_{si} + D'''\bar{v}_n \tag{2.31}$$

and

$$\bar{i}_c = B''''\,\bar{v}_{si} + C''''\,\bar{i}8_{si} + D''''\,\bar{v}_n \tag{2.32}$$

Finally, using the last two equations in (2.28), one has

$$A\bar{v}_n = F_v'\bar{v}_s + F_I'\bar{i}_s + B\bar{v}_n \tag{2.33}$$

Now,

$$(A - B)\bar{v}_n = F_v'\bar{v}_s + F_I'\bar{i}_s \tag{2.34}$$

This equation can be solved for the node voltages, provided that $A - B$ is nonsingular. This is even more problematic than for the case without dependent sources because the matrix B is a function of the gain coefficients of the dependent sources; for some set of such values the solution might exist and for others it might not. In any event if $A - B$ is nonsingular, one obtains once more a response that is linear with respect to the vector of independent source values.

Figure 2.27 shows a rather complex example circuit with dependent sources. As pointed out earlier, there are often reasons for preferring one reference node to another. Here, notice that if we choose one of the nodes to which a voltage source is attached it is not necessary to write a nodal equation for the nonreference node because, when the circuit is tested by deactivation of *all* the sources, the node disappears into the ground reference; thus, it is part of a generalized node including the reference called a **nonessential node**. For this circuit, choose the node at the bottom of the 2*V* independent source. The resulting circuit, prepared for nodal analysis, is shown in Figure 2.28. Surfaces have been drawn around both generalized nodes and the one essential node and they have been shaded for emphasis. Note that we have chosen one node voltage within the one supernode arbitrarily and have expressed the other node

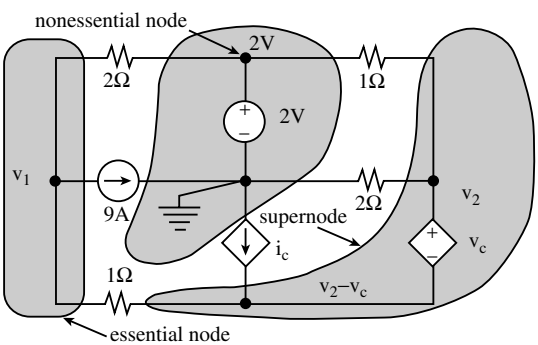


FIGURE 2.28 The example circuit prepared for nodal analysis.

voltage within that supernode in terms of the first and the voltage source value; furthermore, we have taped both dependent sources and written in the known value at the one nonessential node.

The nodal equations for the supernode and for the essential node are

$$\frac{v_1 - 2}{2} + \frac{v_1 - (v_2 - v_c)}{1} = -9 \quad \text{(essential node)} \tag{2.35}$$

and

$$\frac{v_2}{2} + \frac{v_2 - 2}{1} + \frac{v_2 - v_c - v_1}{1} = i_c \quad \text{(supernode)} \tag{2.36}$$

Now, the two dependent sources are untaped and their values expressed in terms of the unknown node voltages and known values using KVL, KCL, and Ohm's law. This results in (referring to the original circuit for the definitions)

$$v_c = -\frac{3}{2} v_2 \tag{2.37}$$

and

$$i_c = 4 - 2v_2 \cdot \tag{2.38}$$

Solving these four equations simultaneously gives $v_1 = -2$ V and $v_2 = 2$ V.

If the circuit under consideration contains op amps, one can first replace each op amp by a VCVS, using the above procedure, and then allow the voltage gain to go to infinity. This is a bit unwieldy, so one often models the op amp in a different way as a circuit element called a **nullor**. This is explored in more detail elsewhere in the book and is not discussed here.

Thus far, this chapter has considered only nondynamic circuits whose independent sources were all constants (DC). If these independent sources are assumed to possess time-varying waveforms, no essential modification ensues. The only difference is that each node voltage, and hence each circuit variable, becomes a time-varying function. If the circuit considered contains capacitors and/or inductors, however, the nodal equations are no longer algebraic; they become differential equations. The method developed above remains applicable, however. We will now show why.

Capacitors and inductors have the *v–i* relationships given in Figure 2.29. The symbols p and $1/p$ are referred to as **operators**, **differential operators**, or **Heaviside operators**. The last term is in honor of Oliver Heaviside, who first used them in circuit analysis. They are defined by

$$p = \frac{d}{dt} \tag{2.39}$$

$$\frac{1}{p} = \int_{-\infty}^{t} (\,)\, da \tag{2.40}$$

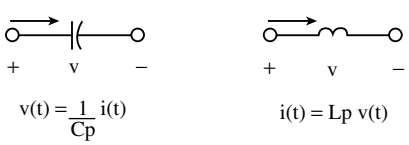


FIGURE 2.29 The dynamic element relationships.

The notation suggests that they are inverses of each other, and this is true; however, one must suitably restrict the signal space in order for this to hold. The most realistic assumption is that the signal space consists of all piecewise continuous functions whose derivatives of all orders exist except on a countable set of points that does not have any finite points of accumulation — *plus all generalized derivatives of such functions*. In fact, Laurent Schwartz, on the first page of the preface of his important work on the theory of distributions, acknowledges that this work was motivated by that of Heaviside. Thus, we will simply assume that all derivatives of all orders of any waveform under consideration exists in a generalized function sense. Higher order differentiation and integration operators are defined in power notation, as expected:

$$p^n = p \cdot p \cdots p = \frac{d^n}{dt^n} \tag{2.41}$$

and

$$\frac{1}{p^n} = \frac{1}{p} \cdot \frac{1}{p} \cdots \frac{1}{p} = \int_{-\infty}^{t} \int_{-\infty}^{b} \cdots \int_{-\infty}^{g} (\,)\, da \tag{2.42}$$

Another fact of the preceding issue often escapes notice, however. Look at any arbitrary function in the above-mentioned signal set, compute its running integral, and differentiate it. This action results in:

$$p\left[\frac{1}{p}x(t)\right] = \frac{d}{dt}\int_{-\infty}^{t} x(\alpha)d\alpha = x(t) \tag{2.43}$$

In fact, it is precisely this property that characterizes the set of all generalized functions. It is closed under differentiation. However, suppose the computation is done in the reverse order:

$$\frac{1}{p}\left[px(t)\right] = \int_{-\infty}^{t} x'(a)\,da = x(t) = x(t) - x(\infty) \tag{2.44}$$

We have assumed here that the Fundamental Theorem of Calculus holds. This is permissible within the framework of generalized functions, provided that the waveform $x(t)$ has a value in the conventional sense at time t. The problem with the previous result is that one does not regain $x(t)$. If it is assumed, however, that $x(t)$ is one sided (that is, $x(t)$ is identically zero for sufficiently large negative values of (t), $x(t)$ will be regained and p and $1/p$ will be inverses of one another. Thus, in the following, we will assume that all independent waveforms are one sided. We will, in fact, interpret this as meaning that they are all zero for $t < 0$. We will also assume that all circuit elements possess one property in addition to their defining *v–i* relationship, namely, that they are causal. Thus, all waveforms in any circuit under consideration will be zero for $t \leq 0$ and the previous two operators are inverses of one another. The only physically reasonable situation in which two-sided waveforms can occur is that of a stable circuit operating in the steady state, which we recognize as being an approximate mode of behavior derived from the previous considerations in the limit as time becomes large.

Referring to Figure 2.29 once more, we define

$$Z_c(p) = \frac{1}{Cp} \tag{2.45}$$

$$Z_L(p) = Lp \tag{2.46}$$

to be the **impedance operators** (or **operator impedances**) for the capacitor and the inductor, respectively. With our one-sidedness causality assumptions, we can manipulate these qualities just as we would manipulate algebraic functions of a real or complex variable.

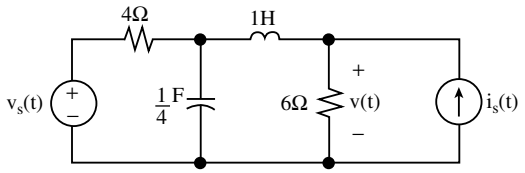


FIGURE 2.30 An example circuit.

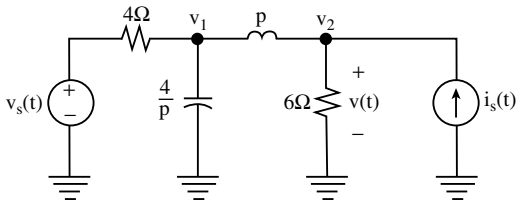


FIGURE 2.31 The example circuit prepared for nodal analysis.

The analysis of a dynamic circuit is illustrated by Figure 2.30. The circuit is shown prepared for nodal analysis, with the reference node at the bottom of the circuit and the dynamic elements expressed in terms of their impedance operators, in Figure 2.31. Note that if the circuit were to contain dependent sources, we would have taped them at this step. The nodal equations at the two essential nodes are

$$\frac{v_1 - v_2}{4} + \frac{v_1}{4/p} + \frac{v_1 - v_2}{p} = 0 \tag{2.47}$$

and

$$\frac{v_2}{6} + \frac{v_2 - v_1}{p} = i_s \tag{2.48}$$

In matrix form, merely rationalizing and collecting terms,

$$\begin{bmatrix} \frac{1}{4} + \frac{p}{4} + \frac{1}{p} & -\frac{1}{p} \\ -\frac{1}{p} & \frac{1}{6} + \frac{1}{p} \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \begin{bmatrix} \frac{1}{4} v_s(t) \\ i_s(t) \end{bmatrix} \tag{2.49}$$

Notice that the coefficient matrix is once again symmetric because no dependent sources exist. Multiplying the first row of each side by $4p$ and the second by $6p$, thus clearing fractions, one obtains

$$\begin{bmatrix} p^2 + p + 4 & -4 \\ -6 & p + 6 \end{bmatrix} \begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \begin{bmatrix} p v_s(t) \\ 6 p i_s(t) \end{bmatrix} \tag{2.50}$$

Now, multiply both sides by the inverse of the 2×2 coefficient matrix to get

$$\begin{bmatrix} v_1(t) \\ v_2(t) \end{bmatrix} = \frac{1}{p(p^2 + 7p + 10)} \begin{bmatrix} p + 6 & 4 \\ 6 & p^2 + p + 4 \end{bmatrix} \begin{bmatrix} p v_s(t) \\ 6 p i_s(t) \end{bmatrix} \tag{2.51}$$

Multiplying the two matrices on the right and cancelling the common p factor (legitimate under our assumptions), we finally have

$$v(t) = v_2(t) = \frac{6v_s(t) + 6\left(p^2 + p + 4\right)i_s(t)}{p^2 + 7p + 10} \tag{2.52}$$

We can, on the one hand, consider the result of our nodal analysis process to be a differential equation which we obtain by cross-multiplication:

$$\left[p^2 + 7p + 10\right]v(t) = 6v_s(t) + 6\left(p^2 + p + 4\right)i_s(t) \tag{2.53}$$

In conventional notation, using the distributive properties of the p operators, one has

$$\frac{d^2v(t)}{dt^2} + 7\frac{dv(t)}{dt} + 10v(t) = 6v_s(t) + 6\frac{d^2i_s(t)}{dt^2} + 6\frac{di_s(t)}{dt} + 24i_s(t). \tag{2.54}$$

On the other hand, it is possible to interpret (2.52) directly as a solution operator equation. We simply note that the denominator factors, then do a partial fraction expansion to get

$$\begin{aligned} v(t) &= \frac{6}{(p+2)(p+5)}v_s(t) + \frac{6\left(p^2 + p + 4\right)}{(p+2)(p+5)}i_s(t) \\ &= \frac{2}{p+2}v_s(t) - \frac{2}{p+2}v_s(t) + 6i_s(t) + \frac{2}{p+2}i_s(t) - \frac{8}{p+2}i_s(t). \end{aligned} \tag{2.55}$$

Thus, we have expressed the two second-order operators in terms of operators of order one. It is quite easy to show that the first-order operator has the following simple form:

$$\frac{1}{p+a}x(t) = e^{-at}\frac{1}{p}\left[e^{at}x(t)\right] \tag{2.56}$$

Using this result, one can quickly show that the impulse and step responses of the first-order operator are

$$h(t) = \frac{1}{p+a}\delta(t) = e^{-at}u(t) \tag{2.57}$$

and

$$s(t) = \frac{1}{p+a}u(t) = \left[1 - e^{-at}\right]u(t) \tag{2.58}$$

respectively. Thus, if $i_s(t) = \delta(t)$ and $v_s(t) = u(t)$, one has

$$v(t) = 6\delta(t) + \frac{1}{5}\left[3 + 5e^{-2t} - 38e^{-5t}\right]u(t) \tag{2.59}$$

References [5, 6] demonstrate that all the usual algebraic results valid for the Laplace transform are also valid for Heaviside operators.

Mesh Analysis

The central concept in nodal analysis is, of course, the node. The central idea in the method we will discuss here is the loop. Just as KCL formed the primary set of equations for nodal analysis, KVL will serve a similar function here. We will begin with the idea of a **mesh**. A mesh is a special kind of loop in

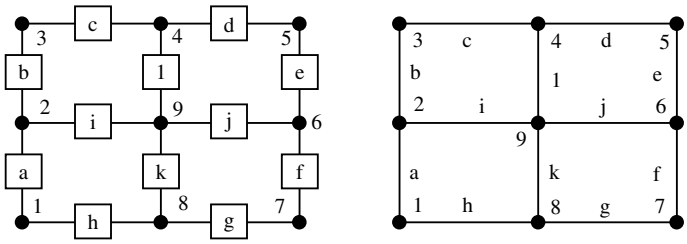


FIGURE 2.32 A circuit and its graph.

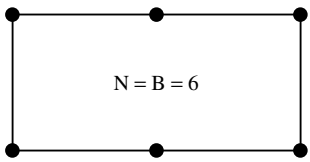


FIGURE 2.33 A one-mesh (series) circuit.

a planar circuit (one that can be drawn on a plane) a loop that does not contain any other loop inside it. If one reflects on this definition a bit, one will see that it depends upon how the circuit is drawn. Figure 2.32 illustrates the idea of a mesh. The nodes have been numbered and the elements labeled with letters for clarity. The circuit graph in Figure 2.32 abstracts all of the information about how the elements are connected, but does not show them explicitly. The lines represent the elements and the solid dots represent the nodes. If we apply the definition given in the introduction to this section, we can quickly verify that {*h, a, i, k*} is a loop. It is a simple loop because each of its elements share only one node with any of the other path elements. It is a mesh because no other loops are inside it.

It is an important fact that the number of meshes in a circuit is given by

$$N_m = B - N + 1 \tag{2.60}$$

where B is the number of branches (elements) and, as usual, N is the number of nodes. To see this, just look at the simple one-mesh graph in Figure 2.33. The number of branches is the same as the number of nodes for such a graph (or circuit). Imagine constructing the graph by placing an element on a planar surface, thereby forming two nodes with the one element. $B - N + 1 = 1 - 2 + 1 = 0$ in this case, and no meshes exist. Now, add another element by connecting one of its leads to one of the leads of the first element. Now, $B - N + 1 = 2 - 3 + 1 = 0$. This can be done indefinitely (or until you tire). At this point, connect one lead of the last element to the free lead of the one immediately preceding and the other lead of the last element to a node already placed. N nodes and $N - 1$ branches will have been put down, and exactly one mesh will have been formed. Thus, it is true that $B - N + 1 = N - (N - 1) + 1 = 1$ mesh and the formula is verified. Now connect a new element to one of the old nodes; the result is that one new element and one new node have been added. A glance at the formula verifies that it remains valid. Again, continue indefinitely, and then connect one new element and no new nodes by connecting the free lead of the last element with one of the nodes in the original one-loop circuit. Clearly, the number of added branches exceeds the number of added nodes by one; once again, the formula is verified. Figure 2.34 shows the new circuit. For the graph shown in the figure, $B = 13$ and $N = 12$, so $B - N + 1 = 2$, as expected. Induction generalizes the result, and (2.60) has been proved.

We now define a fictitious set of currents circulating around the meshes of a circuit. Figure 2.35 illustrates this idea with a circuit graph. All mesh currents are assumed to be circulating in a clockwise direction, although this is not necessary. We see that i_1 is the only current flowing in the branch in which the element current i_a is defined, therefore, $i_a = i_1$; similarly, i_3 is the only mesh current flowing in the element carrying element current i_b, but the two are defined in opposite directions. Thus, one sees that

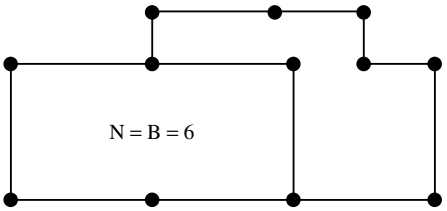


FIGURE 2.34 A two-mesh (series) circuit.

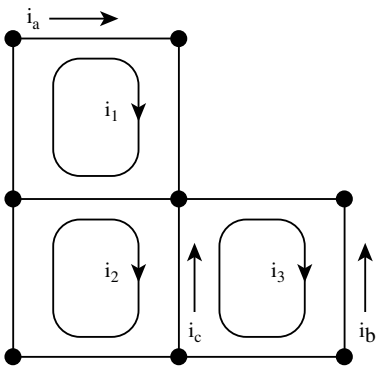


FIGURE 2.35 Mesh currents.

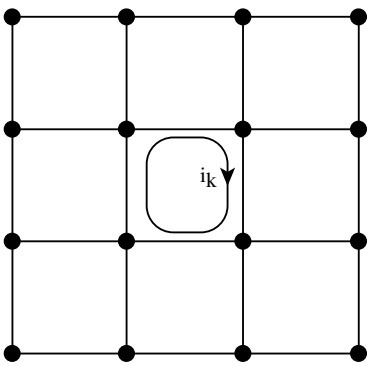


FIGURE 2.36 A fictitious mesh current.

$i_b = -i_3$. The third element where the current is indicated, however, is seen to carry two mesh currents in opposite directions. Hence, its element current is $i_c = i_3 - i_2$. In general, an element that is shared between two meshes has an element current which is the algebraic sum or difference[7] of the two adjacent mesh currents.

We used the term "fictitious" in our definition of mesh current. In the last example, however, we see that it is possible to make a physical measurement of each mesh current because each flows in an element that is not shared with any other mesh. Thus, one need only insert an ammeter in that element to measure the associated mesh current. Circuits exist, however, in which one or more mesh currents are impossible to measure. Figure 2.36 plots the graph of such a circuit. Each of the meshes is assumed to be carrying a mesh current, although only one has been drawn explicitly, i_k. As readily observed, each of the other mesh currents appears in a nonshared branch. For the mesh where the current is shown, however, it is impossible to find an element or a conductor carrying only that current. For this reason, i_k is merely a fiction, though a useful one.

[7]Always the difference if all mesh currents are defined in the same direction: clockwise or counterclockwise.

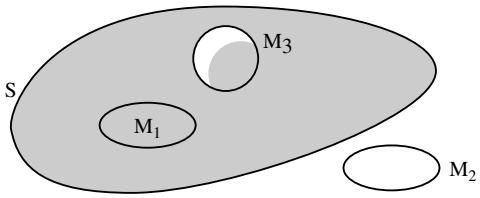


FIGURE 2.37 Illustration of KCL for mesh currents.

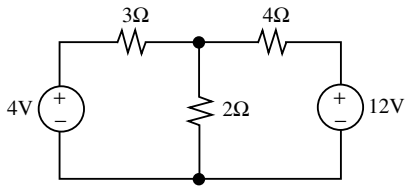


FIGURE 2.38 An example circuit.

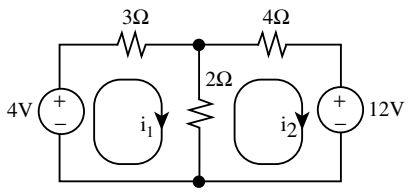


FIGURE 2.39 The example circuit prepared for mesh analysis.

It is easy to see that mesh currents automatically satisfy KCL because they form a complete loop. Observe the stylized picture in Figure 2.37, which shows three meshes represented for simplicity as small, closed ovals. M_1 lies entirely within the arbitrary closed surface S; thus, its current does not appear in KCL for that surface. M_2 lies entirely outside S, so the same thing is true for its current. Finally, we note that, regardless of the shape of S, M_3 penetrates it an even number of times. Thus, its current will appear in KCL for S an even number of times, and half of its appearances will carry a positive sign and half a negative sign. Thus, we have shown that any mesh current automatically satisfies KCL for any closed surface.

Because KCL is automatically satisfied, we must turn to KVL for the solution of a network in terms of its mesh currents. Figure 2.38 is an example circuit. Just as we assumed at the outset of the last subsection that any circuit under consideration contained only resistors and current sources, we will assume at first that any circuit under consideration contains only resistors and voltage sources. The one shown in Figure 2.38 has this property.

The first step is to identify the meshes and assign a mesh current to each. Identification of the meshes is easy, and this is the primary reason for its effectiveness in hand analysis of circuits. The mesh currents can be assigned in arbitrary directions, but for circuits of the sort considered here, it is more convenient to assign them all in the same direction, as in Figure 2.39. Writing one KVL equation for each mesh results in

$$3i_1 + 2(i_1 - i_2) = 4 \tag{2.61}$$

and

$$2(i_2 - i_1) + 4i_2 = -12 \tag{2.62}$$

In matrix form,

$$\begin{bmatrix} 5 & -2 \\ -2 & 6 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} 4 \\ -12 \end{bmatrix} \tag{2.63}$$

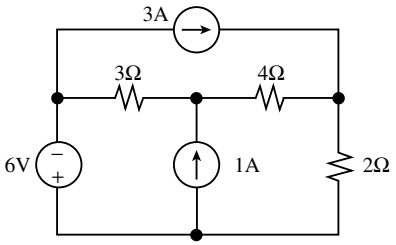


FIGURE 2.40 An example with current sources.

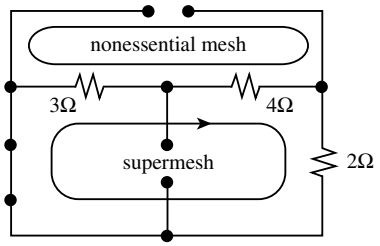


FIGURE 2.41 The deactivated current.

The solution is $i_1 = 0$ A and $i_2 = 2$ A. The same procedure holds for any planar circuit of an arbitrary number of meshes.

Suppose, now, that the circuit being considered has one or more current sources, such as the one in Figure 2.40. The meshes are readily determined; one need only to look for the "window panes", as meshes have been called. The only problem is this: When we write our mesh equations, what values do we use for the voltages across the current sources? These voltages are not known.

Thus, we could ascribe their voltages as unknowns, but this would lead to a hybrid form of analysis in which the unknowns are both element voltages and mesh currents; however, a more straightforward way is available. Consider this question: should the variables we use or the loops around which we decide to write KVL change if we alter the *values* of any of the independent sources? The answer, of course, is no. Thus, let us test the circuit by deactivating it-that is, by reducing all sources to zero. Recalling that a zero-valued voltage source is a short circuit and a zero-valued current source is an open circuit, we obtain the test circuit in Figure 2.41.

Notice what has happened. The two bottom meshes merge, thus forming one larger mesh in the deactivated circuit. The top mesh disappears (as a mesh or loop). For this reason, we refer to the former as a **supermesh** and the latter as a **nonessential mesh**. Observe also that it was the deactivation of the current sources that altered the topology; in fact, deactivation of the voltage sources has no effect on the mesh structure at all. Thus, we see that only one KVL equation is required to solve the deactivated circuit. (Reactivation of the source(s) is necessary, otherwise all voltage and currents will have zero values.) The conclusion relative to our example circuit is this: To solve the original circuit in terms of mesh currents, only one equation (KVL around the supermesh) is necessary.

The original circuit, with its three mesh currents arbitrarily defined, is redrawn in Figure 2.42. Notice that the isolated (nonshared) current source in the top (nonessential) mesh defines the associated mesh current as having the same value as the source itself. On the other hand, the 1 A current source shared by the bottom two meshes introduces a more general constraint: the difference between the two mesh currents must be the same as the source current. This constraint has been used to label the mesh current in the right-hand mesh with a value such that it, minus the left-hand mesh current, is equal to the source current. The nice feature of this approach is that one can clearly see which mesh currents are unknown and which are dependent upon the unknowns. Exactly one independent mesh current is always associated with a supermesh. Recalling our test circuit in Figure 2.41, we see that we need to write only one KVL equation around the supermesh. It is

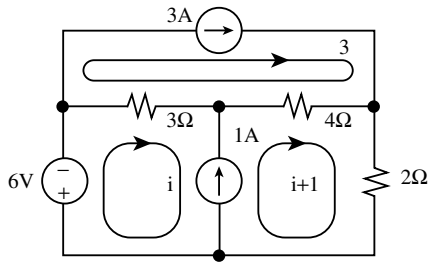


FIGURE 2.42 Assigning the mesh currents.

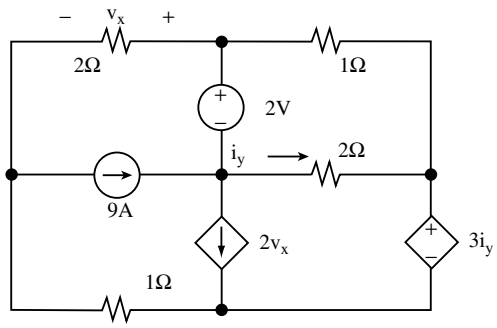


FIGURE 2.43 An example circuit.

$$3(i-3)+4(i+1-3)+2(i+1)=-6 \tag{2.64}$$

or

$$9i-9-8+2=-6 \tag{2.65}$$

The solution is $i = 1$ A. From this, one can compute the mesh current on the bottom right to be $i + 1 = 2$ A and the one in the top loop is already known to be 3 A. With these known mesh currents, we can solve for any circuit variable desired.

The development of mesh analysis seems at first glance to be the complete analog of nodal. This is not quite the case, however, because nodal will work for nonplanar circuits, while mesh works only for planar circuits; furthermore, no global reference exists for mesh currents as it does for node voltages.

Analyzing the problem, we observe that each current source, when deactivated, reduces the number of meshes by one. (A given element can be shared only by two meshes). Combining this fact with (2.60), we see that the required number of mesh equations is

$$N_{me} = B - N + 1 - N_I, \tag{2.66}$$

where, as usual, B is the number of branches, N is the number of nodes, and (in this equation) N_I is the number of current sources.

Note that mesh analysis is undertaken for circuits containing dependent sources in exactly the same manner as in nodal analysis — that is, by first taping the dependent sources, writing the mesh equations as above, and then untaping the dependent sources and expressing their controlled variables in terms of the unknown mesh currents. Figure 2.43 shows an example of such a circuit; in fact, it is the same figure investigated with nodal analysis in the preceding subsection.

The first step is to tape the dependent sources, thus placing them on the same footing as their more independent relatives. Then the circuit is tested for complexity by deactivating it as shown in Figure 2.44. All element labels have been removed merely to avoid obscuring the ideas being discussed. We see one supermesh, one essential mesh, and no nonessential mesh. Therefore, two KVL equations must be written in the original circuit, which is shown with the dependent sources taped in Figure 2.45. Notice that only

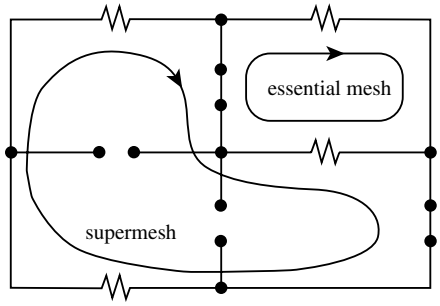


FIGURE 2.44 Testing the example circuit.

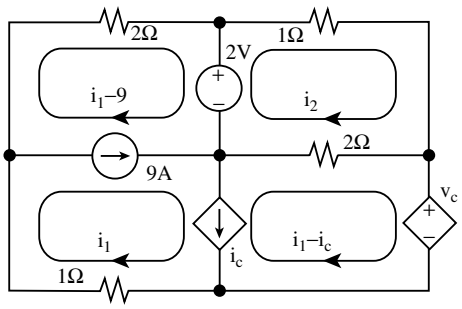


FIGURE 2.45 The example circuit prepared for mesh analysis.

two unknowns exist, and also that the dependent sources have been taped. For the moment, we have turned them into independent sources (albeit with unknown values).

We are now in a position to write KVL equations:

$$2(i_1 - 9) + 2(i_1 - i_c - i_2) + 1i_1 = -2 - v_c \qquad \text{(supermesh)} \tag{2.67}$$

and

$$1i_2 + 2(i_2 - i_1 + i_c) = 2 \qquad \text{(essential mesh)} \tag{2.68}$$

Observe that i_c and v_c are not known quantities, as would be the case were they the values of independent sources. Thus, at this point, we must untape the dependent sources and express their values in terms of the mesh currents. We find that

$$i_c = 2v_x = 2x2x(-i_1 + 9) = -4i_1 + 36 \tag{2.69}$$

and

$$v_c = 3i_y = 3(i_1 - i_c - i_2) = 15i_1 - 3i_2 + 36 \tag{2.70}$$

Inserting the last two results in (2.67) and (2.68) results in the matrix equation

$$\begin{bmatrix} 28 & -5 \\ -10 & 3 \end{bmatrix} \begin{bmatrix} i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} 196 \\ -70 \end{bmatrix} \tag{2.71}$$

The coefficient matrix is no longer symmetric now that dependent sources have been introduced. (This is also the case with nodal analysis. The example treating this same circuit is found in the last subsection and should be checked to verify this point.) However, the solution is found, as usual, to be $i_1 = 7$ A and $i_2 = 0$ A.

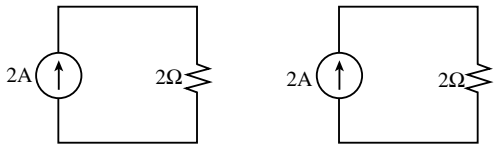


FIGURE 2.46 An example circuit.

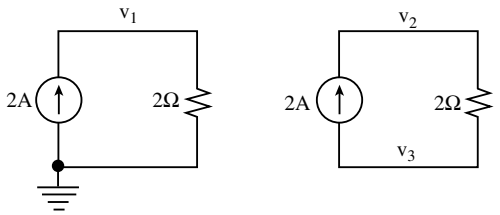


FIGURE 2.47 The example circuit prepared for nodal analysis.

A careful consideration of what we have done up to this point reveals that the mesh equations can be written in the form

$$A\bar{i}_M = B\bar{i}_s + C\bar{v}_s \tag{2.72}$$

In this general formulation, A is a square $n_M \times n_M$ matrix, where m is the number of meshes, and B and C are rectangular matrices whose dimensions depend upon the number of independent voltage and current sources, respectively. The variables $\bar{v}_s$ and $\bar{i}_s$ are the column matrices of independent voltage and current source values, respectively. A is symmetric if the circuit contains only resistors and independent sources. As was the case for nodal analysis, the elucidation of conditions under which the A matrix is nonsingular is difficult. Certainly, it can be singular for circuits with dependent sources; surprisingly, circuits also exist with only resistors and independent sources for which A is singular as well.

Finally, the mesh analysis procedure for circuits with dynamic elements should be clear. The algebraic process closely follows that for nodal analysis. For this reason, that topic is not discussed here.

Fundamental Cutset-Loop Circuit Analysis

As effective as nodal and mesh analysis are in treating circuits by hand, particular circuits exist for which they fail. Consider, for example, the circuit in Figure 2.46. If we were to blindly perform nodal analysis on this circuit, we would perhaps prepare it for analysis as shown in Figure 2.47. We have three nonreference nodes, hence, we have three nodal equations:

$$\frac{v_1}{2} = 2 \tag{2.73}$$

$$\frac{v_2 - v_3}{2} = 2 \tag{2.74}$$

$$\frac{v_3 - v_2}{2} = -2 \tag{2.75}$$

The third equation is simply the negative of the second; hence, the set of equations is linearly dependent and does not have a unique solution. The reason is quite obvious: the circuit is not connected.[8] It actually consists of two circuits considered as one. Therefore, one should actually select two reference nodes rather than one. A bit of reasoning along this line indicates that the number of nodal equations should be

[8]A circuit is connected if at least one path exists between each pair of nodes.

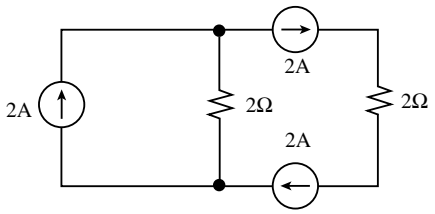


FIGURE 2.48 Another example circuit.

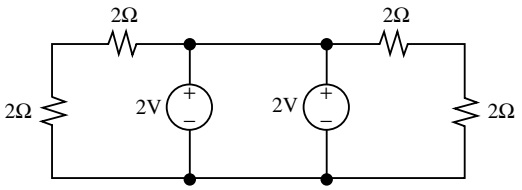


FIGURE 2.49 Another example of a singular circuit.

$$N_{ne} = N - 1 - P \tag{2.76}$$

where P is the number of separate parts, and hence the number of reference nodes required.

Another way nodal analysis can fail is not quite as obvious. Figure 2.48 illustrates the situation. In this case, we have a cutset of current sources. Therefore, in reality, at least one of the current sources cannot be independent for it must have the same value as the other in the cutset. We will leave the writing of the nodal equations as an exercise for the reader. They are, however, linearly dependent. The problem here clearly becomes evident if one deactivates the circuit, because the resulting test circuit is not connected.

Analogous problems can occur with mesh analysis, as the circuit in Figure 2.49 demonstrates. We find a loop of voltage sources and, when the circuit is deactivated, one mesh disappears. Again, it is left as an exercise for the reader to write the mesh equations and show that they are linearly independent (the coefficients of all currents in the KVL equation for the central mesh are zero).

One might question the practically of such circuits because clearly no one would design such networks to perform a useful function. In the computer automation of circuit analysis, however, dynamic elements are often modeled over a small time increment in terms of independent sources, and singular behavior can result. Furthermore, one would like to be able to include more general elements than R, L, C, and voltage and current sources. For such reasons, a general method that does not fail is desirable. We develop this method next. It is related to the modified nodal analysis technique that is described elsewhere in the book.

To develop this technique, we will examine the graph of a specific circuit: the one in Figure 2.50. Graph theory is covered elsewhere in the book, but salient points will be reviewed here [7, 8]. The graph, of course, is not concerned at all with the *v–i* characteristics of the elements themselves — only with how they are interconnected. The lines (or edges or branches) represent the elements and the dots represent the nodes. The arrows represent the assumed references for voltage and current, the positive voltage at the "upstream" end of the arrow and the current reference in the direction of the arrow. We also recall the definition of a tree: for a connected graph of N nodes, a **tree** is any subset of edges of the graph that connects all the nodes, but which contains no loops. Such a tree is shown by means of the darkened edges in the figure: *a, b,* and *c.* The complement of a tree is called a **cotree**. Thus, edges *d, e, f, g,* and *h* form a cotree in Figure 2.50. If a graph consists of separate parts (that is, it is not connected), then one calls a subset of edges that connects all N nodes, but forms no loops, a **forest**. The complement of a forest is a **coforest**. The analysis method presented here is applicable to either connected or nonconnected circuits. However, we will use the terms for a graph that is connected for ease of comprehension; one

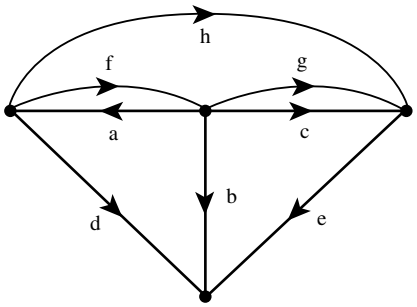


FIGURE 2.50 An example of a circuit graph.

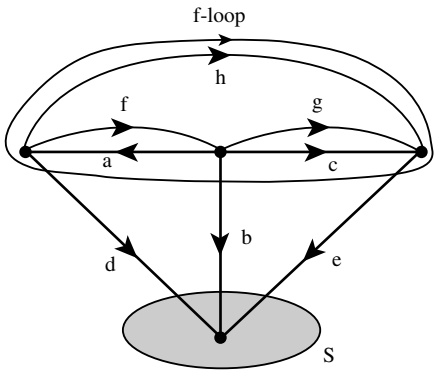


FIGURE 2.51 Fundamental cutsets and loops.

should go through a parallel development for nonconnected circuits to assure oneself that the generalization holds.

Each edge contained in a tree is called a **twig**, and each edge contained in a cotree is called a **link**. (Remember that the set of all nodes in a connected graph is split into exactly two sets of nodes, each of which is individually connected by twigs, when a tree edge is removed). The set of links having one of its nodes in one set and another in the second, together with the associated twig defining the two sets of nodes, is called a **fundamental cutset**, or ***f*-cutset**. If all links associated with a given tree are removed, then the links replaced one at a time, it can be seen that each link defines a loop called a **fundamental loop** or ***f*-loop**. Figure 2.51 is a fundamental cutset and a fundamental loop for the graph shown in Figure 2.50. The closed surface *S* is placed around one of the two sets of nodes so defined (in this case consisting of a single node) and is penetrated by the edges in the cutset *b*, *d*, and *e*. A natural orientation of the cutset is provided by the direction of the defining twig, in this case edge *b*. Thus, a positive sign is assigned to *b*; then, any link in the fundamental cutset with a direction relative to *S* agrees with that of the twig receives a positive sign, and each with a direction that is opposite receives a negative sign. Similarly, the fundamental loop is given a positive sense by the direction of the defining link, in the case edge *h*. It is assigned a positive sign; then, each twig in the *f*-loop is given a positive sign if its direction coincides in the loop with the defining link, and a negative sign if it does not.

The positive and negative signs just defined can be used to write one KCL equation for each *f*-cutset and one KVL equation for each *f*-loop, as follows. Consider the example graph with which we are working. The *f*-cutsets are $\{d, b, e\}$, $\{d, a, f, h\}$, and $\{e, c, g, h\}$. In general, $N - 1$ *f*-cutsets are associated with each tree — exactly the same as the number of twigs in the tree. Using surfaces similar to *S* in Figure 2.51 for each of the *f*-cutsets, we have the following set of KCL equations:

$$i_a - i_d - i_f - i_h = 0 \tag{2.77}$$

$$i_b + i_d + i_e = 0 \tag{2.78}$$

$$i_c - i_e + i_g + i_h = 0 \tag{2.79}$$

In matrix form, these equations become

$$\begin{bmatrix} 1 & 0 & 0 & -1 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 1 & 1 \end{bmatrix} \begin{bmatrix} i_a \\ i_b \\ i_c \\ i_d \\ i_e \\ i_f \\ i_g \\ i_h \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \end{bmatrix}. \tag{2.80}$$

The coefficient matrix consists of zeroes, and positive and negative ones. It is called the ***f*-cutset matrix**. Each row corresponds to the KCL equation for one of the *f*-cutsets, and has a zero entry for each edge not in that cutset, $a + 1$ for any edge in the cutset with the same orientation as the defining twig, and $a - 1$ for each edge in the cutset whose orientation is opposite to the defining twig, and $a - 1$ for each edge in the cutset whose orientation is opposite to the defining twig. One often labels the rows and columns,

$$Q = \begin{matrix} \\ a \\ b \\ c \end{matrix} \begin{bmatrix} a & b & c & d & e & f & g & h \\ 1 & 0 & 0 & -1 & 0 & -1 & 0 & -1 \\ 0 & 1 & 0 & 1 & 1 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & -1 & 0 & 1 & 1 \end{bmatrix}. \tag{2.81}$$

to emphasize the relation between the matrix and the graph. Thus, the first row corresponds to KCL for the *f*-cutset defined by twig *a*, the second to that defined by twig *b*, and the last to the cutset defined by twig *c*. The columns correspond to each of the edges in the graph, with the twigs occupying the first $N - 1$ columns in the same order as that in which they appear in the rows. Notice that a unit matrix of order $N - 1 \times N - 1$ is located in the leftmost $N - 1$ columns. Furthermore, Q has dimensions $(N - 1) \times B$, where B is the number of branches (edges). Clearly, Q has maximum rank because of the leading unit matrix. More succinctly, one writes KCL in terms of the *f*-cutset matrix as

$$Q\bar{i} = \left[U \vdots H\right]\bar{i} = 0, \tag{2.82}$$

where $\bar{i}$ is the column matrix of all the branch currents. Here, the structure of Q appears explicitly with the unit matrix in the first $N - 1$ columns and, for our example,

$$H = \begin{bmatrix} -1 & 0 & -1 & 0 & -1 \\ 1 & 1 & 0 & 0 & 0 \\ 0 & -1 & 0 & 1 & 1 \end{bmatrix} \tag{2.83}$$

In general, H will have dimensions $(N - 1) \times (b - N + 1)$.

Each of the links, all $B - N + 1$ of them, have an associated KVL equation. In our example, using the same order for these links and equations that occurs in the KCL equations,

$$\begin{bmatrix} 1 & -1 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 0 & 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\ 0 & 0 & -1 & 0 & 0 & 0 & 1 & 0 \\ 1 & 0 & -1 & 0 & 0 & 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} v_a \\ v_b \\ v_c \\ v_d \\ v_e \\ v_f \\ v_g \\ v_h \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ 0 \\ 0 \\ 0 \end{bmatrix} \tag{2.84}$$

We have one row for each link and, therefore, one for each *f*-loop. If a given edge is in the given loop, $a + 1$ is in the corresponding column if its direction agrees with that of the defining link, and $a - 1$ if it disagrees. Notice that a unit matrix of dimensions $(B - N + 1) \times (B - N + 1)$ is located in the last $B - N + 1$ columns. Even more important, observe that the matrix in the first $N - 1$ columns has a familiar form; in fact, it is $-H'$, the negative transpose of the matrix in (2.83).

This is no accident. In fact, the entries in this matrix are strictly due to twigs in the tree. Focus on a given twig and a given link. The twig defines two twig-connected sets of nodes, as mentioned above. If the given link has both its terminal nodes in only one of these sets, the given twig voltage does not appear in the KVL equation for that *f*-loop. If, on the other hand, one of the link nodes is in one of those sets and the other in the alternate set, the given twig voltage will appear in the KVL equation for the given link, with $a + 1$ multiplier if the directions of the twig agree relative to the *f*-loop and $a - 1$ if they do not. However, a little thought shows that the same result holds for the *f*-cutset equation defined by the twig, except that the signs are reversed. If the link and twig directions agree for the *f*-loop, they disagree for the *f*-cutset, and vice versa. Thus, we can write KVL for the *f*-loops, in general, as

$$B_f \bar{v} = \left[-H' \vdots U\right] \bar{v} = 0 \tag{2.85}$$

B_f is called the fundamental loop matrix, and $\bar{v}$ is the column matrix of all branch voltages.

Suppose that we partition the branch voltages and branch currents according to whether they are associated with twigs or links. Thus, we write

$$\bar{v} = \left[\bar{v}_T' \; \bar{v}_C'\right]' \tag{2.86}$$

and

$$\bar{i} = \left[\bar{i}_T' \; \bar{i}_C'\right]'. \tag{2.87}$$

We use transpose notation to conserve space, and the subscripts *T* and *C* represent tree and cotree voltages and currents, respectively. We cannot use (2.82) and (2.85) to write the composite circuit variable vector as

$$\bar{u} = \begin{bmatrix} \bar{v} \\ \bar{i} \end{bmatrix} = \begin{bmatrix} U & 0 \\ H' & 0 \\ 0 & -H \\ 0 & U \end{bmatrix} \begin{bmatrix} \bar{v}_T \\ \bar{i}_C \end{bmatrix} \tag{2.88}$$

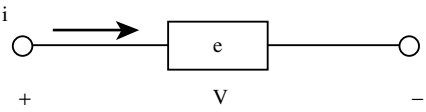


FIGURE 2.52 A two-terminal element.

The coefficient matrix is of dimensions $2B \times B$ and has rank B because of the two unit matrices. What we have accomplished is a direct sum decomposition of the $2B$-dimensional vector space consisting of all circuit variables in terms of the $N - 1$ dimensional vector space of tree voltages and the $B - N + 1$ dimensional vector space of link currents. Furthermore, the tree voltages and link currents form a basis for the vector space of all circuit variables.

We have discussed topology enough for our purposes. We now treat the elements. We are looking for a generalized method of circuit analysis that will succeed, not only for circuits containing *R, L, C,* and source elements, but ideal transformers, gyrators, nullators, and norators (among others) as well. Thus, we turn to a discussion of elements. [9].

Consider elements having two terminals only, as shown in Figure 2.52. The most general assumption we can make, assuming that we are ruling out "nonlinear" elements, is that the *v–i* characteristic of each such element be *affine;* that is, it is defined by an operator equation of the form

$$\begin{bmatrix} a & b \\ 0 & c \end{bmatrix} \begin{bmatrix} v \\ \bar{i} \end{bmatrix} = \begin{bmatrix} f(t) \\ g(t) \end{bmatrix} \tag{2.89}$$

where the parameters a, b, and c are operators. It is more classical to assume a scalar form of this equation; that is, with c and $g(t)$ both zero. In a series of papers in the 1960s, however, Carlin and Youla, Belevitch, and Tellegen [10–12] proposed that the *v–i* characteristic be interpreted as a multidimensional relationship. Among other things to come out of the approach was the definition of the nullator and the norator. Now, assuming that this defining characteristic is indeed multidimensional, we see at once that it is not necessary to consider operator matrices of a dimension larger than 2×2. There must be two columns because there are only two scalar terminal variables. If more than two rows were found, the additional equations would be either redundant or inconsistent, depending upon whether row reductions resulted in additional rows of all zeroes or in an inconsistent equation. Finally, the (2, 1) element in the operator matrix clearly can be chosen to be zero as shown, because otherwise it could be reduced to zero with elementary row operations. That is, one could, unless $a = 0$; but here, an exchange of rows produces the desired result shown. Note that any or all of *a, b,* and *c* can be the zero operator.

We pause here to remark that *a* and *b* can be rather general operators. If they are differential, or Heaviside, operators, (that is, they are real, rational functions of p), a theory of lumped circuits (differential systems) is obtained. On the other hand, they could be rational functions of the delay operator E.[9] In this case, one would obtain the theory of distributed (transmission line) circuits. Then, if a common delay parameter is used, one obtains a theory of commensurate transmission line circuits; if not, an incommensurate theory results. If the parameters are functions of both p and d, a mixed theory results. We will assume here that *a, b*, and *c* are rational functions of p.

Let us suppose that c is the zero operator and that $g(t) = 0$ is the second equality resulting from the stipulation of existence (consistency). This gives the affine scalar relationship

$$av + bi = f(t) \tag{2.90}$$

[9] $Ex(t) = x(t - T)$ for all t and all waveforms $x(t)$.

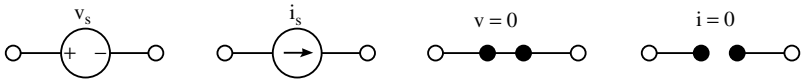


FIGURE 2.53 Conventional two-terminal elements.

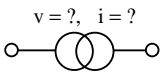


FIGURE 2.54 The norator.

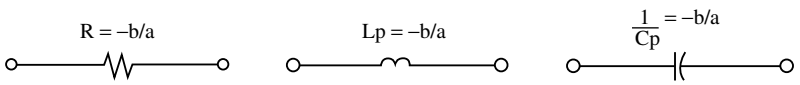


FIGURE 2.55 Passive elements.

Special cases are now examined. For instance, if $b = 0$ and $a \neq 0$, one has

$$v(t) = f(t)/a = v_s(t) \tag{2.91}$$

This, of course, is the *v–i* characteristic for an independent voltage source. If, on the other hand, $a = 0$ and $b \neq 0$, one has

$$i(t) = f(t)/b = i_s(t) \tag{2.92}$$

This is an ideal current source. If, in addition $f(t)$ is identically zero, one obtains a short circuit and an open circuit, respectively. These results are shown in Figure 2.53. Now suppose that a and b are both zero. Then, $f(t)$ must be identically zero as well; otherwise, the element does not exist. In this case any arbitrary voltage and current are possible. The resulting element, a "singular one" to be sure, is called a **norator**. Its symbol is shown in Figure 2.54.

Remaining with the same general case, that is, with $c = 0$ and $g(t) = 0$, we ask what element results if we also assume that neither a nor b are zero, but that $f(t)$ is identically zero. We can solve for either the voltage or the current. In either case, one obtains a passive element, as shown in Figure 2.55. If $-b/a$ is constant, a resistor will result; if $-b/a$ is a constant times the differential operator p, an inductor will result; and if $-b/a$ is reciprocal in p, a capacitor will result. In case the ratio is a more complicated function of p, one would consider the two-terminal object to be a subcircuit, that is, a two-terminal object decomposable into other elements, with $-b/a$ being the driving point impedance operator.

One can, in fact, derive the Thévenin and Norton equivalents from these considerations. Staying with the general case of c and g (t) both zero, but allowing $f(t)$ to be nonzero, we first assume that $a \neq 0$. Then, we obtain

$$v(t) = \frac{f(t)}{a} - \frac{b}{a} i(t) = v_{oc}(t) + Z_{eq}(p) i(t) \tag{2.93}$$

which represents the Thévenin equivalent subcircuit shown in Figure 2.56a. Alternately, if $b \neq 0$ we can write

$$i(t) = \frac{f(t)}{b} - \frac{a}{b} i(t) = i_{sc}(t) + Y_{eq}(p) v(t) \tag{2.94}$$

The latter equation is descriptive of the Norton equivalent shown in Figure 2.56b. The basic assumption is that the two-terminal object has a *v–i* characteristic (i.e., an affine relationship); if this object contains

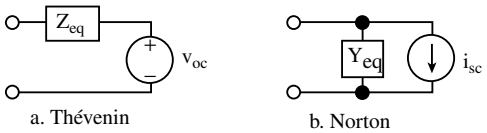


FIGURE 2.56 Two general equivalent subcircuits.

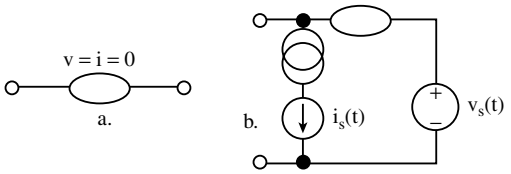


FIGURE 2.57 The nullator element and an equivalent subcircuit.

only elements characterized by affine relationships having a rank property to be given later, one can use the analysis method being presented here to prove that this assumption is true. At this point, however, we are merely assembling a catalog of elements, so we assume that the two-terminal object is, indeed, a single element (it cannot be decomposed farther).

We have only one other case to consider: that in which c is a nonzero operator. If this is the situation and if, in addition, a $\neq 0$ as well, one can solve (2.90) by inverting the coefficient matrix to obtain

$$\begin{bmatrix} v \\ i \end{bmatrix} = \begin{bmatrix} 1/c & -b/ac \\ 0 & 1/a \end{bmatrix} \begin{bmatrix} f(t) \\ g(t) \end{bmatrix} = \begin{bmatrix} v_s(t) \\ i_s(t) \end{bmatrix} \tag{2.95}$$

Therefore, the voltage and current are independently specified. First, suppose that both $v_s(t)$ and $i_s(t)$ are identically zero. Then, one has $v(t) = 0$ and $i(t) = 0$ for t. The associated element is called a **nullator**, and has the symbol shown in Figure 2.57(a). Finally, if $v_s(t)$ and $i_s(t)$ are nonzero, one can sketch the equivalent subcircuit as in Figure 2.57(b).

At this point, we have an exhaustive catalog of two-terminal circuit elements: the independent voltage and current sources, the resistor, the inductor, the capacitor, the norator, and the nullator. We would like to include more complex elements with more than two terminals as well. Figure 2.58(a) shows a three-terminal element and Figure 2.58(b) shows a two-port element. For the former, we see at once that only two voltages and two currents can be independently specified because KVL gives the voltage between the left and right terminals in terms of the two shown, while KCL gives the current in the third lead. As for the latter, it is a basic assumption that only the two-port voltages and the two-port currents are required to specify its operation. In fact, one assumes that the currents coming out of the bottom leads are identical to those going into the top leads. We also assume that the v–i characteristic is independent of the voltages between terminals in different ports. Each of the ports will be an edge in the circuit graph that results when such elements are interconnected.

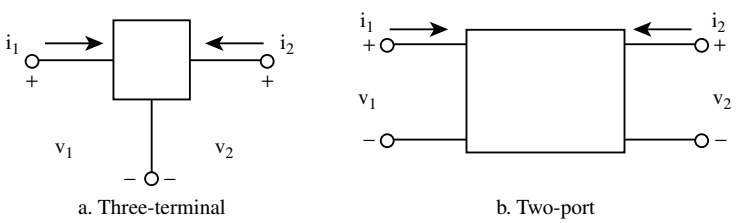


FIGURE 2.58 Three-terminal and two-port elements.

Because four variables are associated with a three-terminal or two-port element, the dimensionality of the resulting vector space is four; thus, we assume that the describing *v–i* characteristic is

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \\ 0 & 0 & a_{33} & a_{34} \\ 0 & 0 & 0 & a_{44} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} f_1(f) \\ f_2(f) \\ f_3(f) \\ f_4(f) \end{bmatrix} \tag{2.96}$$

We justify this form exactly as for the case of two-terminal elements. We will not exhaustively catalog all of the possible three-terminal/two-port elements for reasons of space; however, note that the usual case is that in which $a_{ij} = 0$ for $i \geq 3$. In this case one must insist that $f_3(t) = f_4(t) = 0$; then one has the 2×2 system of equations

$$\begin{bmatrix} a_{11} & a_{12} & a_{13} & a_{14} \\ 0 & a_{22} & a_{23} & a_{24} \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} f_1(t) \\ f_2(t) \end{bmatrix} \tag{2.97}$$

If the two forcing functions on the right are not identically zero, a number of different two-port equivalent circuits can be generated — generalized Thévenin and Norton equivalents. If both of these forcing functions are identically zero and if at least one 2×2 submatrix of the coefficient operator matrix on the left side is nonsingular, one can derive a **hybrid matrix** and a **hybrid parameter equivalent circuit**. Specialized versions are the impedance parameters, the admittance parameters, and the transmission or chain parameters. Furthermore, one can accommodate controlled sources, transformers, gyrators, and all of the other known two-port elements.

To present just one example, assume that the operator (2.97) has the form

$$\begin{bmatrix} n & -1 & 0 & 0 \\ 0 & 0 & -1 & n \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.98}$$

The parameter n, assumed to be a real scalar multiplier, is called the **turns ratio**, and the element is the ideal transformer. The VCVS (voltage controlled voltage source) obeys

$$\begin{bmatrix} \mu & -1 & 0 & 0 \\ 0 & 0 & 1 & 0 \end{bmatrix} \begin{bmatrix} v_1 \\ v_2 \\ i_1 \\ i_2 \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix} \tag{2.99}$$

Thus, i_1 is identically zero and $v_2 = \mu v_1$. The quantity μ is the voltage gain.

Similarly, for each element with any number of ports,[10] we can write

[10]A two-terminal element is a one-port device.

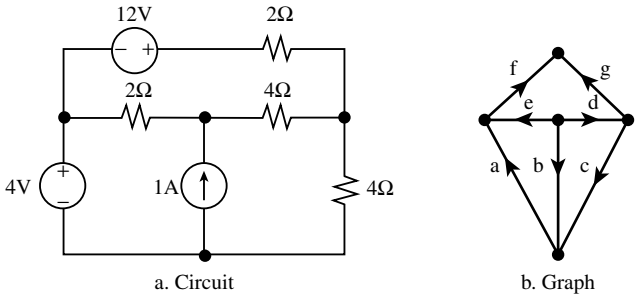


FIGURE 2.59 An example circuit and its graph.

$$A_0\bar{v} + B_0\bar{i} = \bar{C}_0 \tag{2.100}$$

where the voltage and current vectors are the terminal variables of the element. We can then represent the element equations for any circuit in the same form by forming A_0 and B_0 as quasidiagonal matrices, each of whose diagonal terms is the corresponding A_0 or B_0 for a given element, and stacking up the individual $\bar{C}_0$ column matrices to form the overall matrix. We then rewrite (2.100) in the form

$$\left[A_0 B_0\right]\begin{bmatrix}\bar{v}\\ \bar{i}\end{bmatrix} = \bar{C} \tag{2.101}$$

where the voltage and current vectors are each $B \times 1$ column matrices of the individual element voltages and currents. *We make the assumption that the matrix* $[A\ B]$, *which is of dimension* $B \times 2B$ *is of maximum rank b.* This is the only assumption required for the procedure to be outlined to succeed, as will later be demonstrated.

An example will clarify things. Figure 2.59 is an example circuit. The correspondence between the edge labels and the circuit elements is obvious; that is, for instance, *a* is the 4 V voltage source and its voltage is –4 V (minus, because of the definition of positive voltage on edge *a* in the graph). We have shown a tree on the graph. The *f*-cutset matrix is

$$Q = \begin{array}{c} \\ b \\ d \\ e \\ f \end{array}\begin{bmatrix} b & d & e & f & a & c & g \\ 1 & 0 & 0 & 0 & -1 & 1 & 0 \\ 0 & 1 & 0 & 0 & 0 & -1 & -1 \\ 0 & 0 & 1 & 0 & 1 & 0 & 1 \\ 0 & 0 & 0 & 1 & 0 & 0 & 1 \end{bmatrix} = \left[U \vdots H\right] \tag{2.102}$$

Thus,

$$H = \begin{bmatrix} -1 & 1 & 0 \\ 0 & -1 & -1 \\ 1 & 0 & 1 \\ 0 & 0 & 1 \end{bmatrix} \tag{2.103}$$

Although we could construct it from the *Q* matrix, we can just as easily read off the *f*-loop matrix from the graph:

$$B = \begin{matrix} a \\ c \\ g \end{matrix} \begin{bmatrix} b & d & e & f & a & c & g \\ 1 & 0 & -1 & 0 & 1 & 0 & 0 \\ -1 & 1 & 0 & 0 & 0 & 1 & 0 \\ 0 & 1 & -1 & -1 & 0 & 0 & 1 \end{bmatrix} = \left[-H' \vdots U\right] \tag{2.104}$$

The element constraint equations are

$$\begin{bmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -2 \end{bmatrix} \begin{bmatrix} v_b \\ v_d \\ v_e \\ v_f \\ v_a \\ v_c \\ v_g \\ i_b \\ i_d \\ i_e \\ i_f \\ i_a \\ i_c \\ i_g \end{bmatrix} = \begin{bmatrix} -1 \\ 0 \\ 0 \\ -12 \\ 4 \\ 0 \\ 0 \end{bmatrix} \tag{2.105}$$

In this case, both A_0 and B_0 are diagonal because all the elements are of the two-terminal variety.

The vector of all circuit variables is now expressed in terms of the basis in (2.88), the tree voltages and link currents. We then have

$$\left[AB\right] \begin{bmatrix} \bar{v} \\ \bar{i} \end{bmatrix} = \left[AB\right] \begin{bmatrix} U & 0 \\ H' & 0 \\ 0 & -H \\ 0 & U \end{bmatrix} \begin{bmatrix} \bar{v}_T \\ \bar{i}_C \end{bmatrix}$$

$$
= \begin{bmatrix}
0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -2 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -4 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & -2
\end{bmatrix}
$$

$$
\times \begin{bmatrix}
1 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 1 & 1 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & -1 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 1 \\
0 & 0 & 0 & 0 & -1 & 0 & -1 \\
0 & 0 & 0 & 0 & 0 & 0 & 1 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 1 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 1
\end{bmatrix}
\begin{bmatrix} v_b \\ v_d \\ v_e \\ v_f \\ v_a \\ i_c \\ i_g \end{bmatrix}
= \begin{bmatrix} -1 \\ 0 \\ 0 \\ -12 \\ 4 \\ 0 \\ 0 \end{bmatrix} \tag{2.106}
$$

After multiplying the two matrices, we have a more compact matrix

$$
\begin{bmatrix}
0 & 0 & 0 & 0 & 1 & -1 & 0 \\
0 & 1 & 0 & 0 & 0 & -4 & -4 \\
0 & 0 & 1 & 0 & 2 & 0 & 2 \\
0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-1 & 0 & 1 & 0 & 0 & 0 & 0 \\
1 & -1 & 0 & 0 & 0 & -4 & 0 \\
0 & -1 & 1 & 1 & 0 & 0 & -2
\end{bmatrix}
\begin{bmatrix} v_b \\ v_d \\ v_e \\ v_f \\ i_a \\ i_c \\ i_g \end{bmatrix}
= \begin{bmatrix} -1 \\ 0 \\ 0 \\ -12 \\ 4 \\ 0 \\ 0 \end{bmatrix} \tag{2.107}
$$

We leave it to the reader to show that the solution is given by (in transpose notation):

$$\begin{bmatrix} v_b & v_d & v_e & v_f & i_a & i_c & i_g \end{bmatrix}' = \begin{bmatrix} 0 & -4 & 4 & -12 & 0 & 1 & -2 \end{bmatrix} \tag{2.108}$$

In general, one must solve the matrix equation

$$\begin{bmatrix} A\ B \end{bmatrix} = \begin{bmatrix} U & 0 \\ H' & 0 \\ 0 & -H \\ 0 & U \end{bmatrix} \begin{bmatrix} \bar{v}_T \\ \bar{i}_C \end{bmatrix} = C \tag{2.109}$$

where H is the nonunit submatrix in the f-cutset matrix and C_0 is the $B \times 1$ column matrix of constants (or independent functions of time). Here is the crucial result: $[A_0\ B_0]$ has dimensions $B \times 2B$; if it has rank B, then the product of it with the next matrix, which also has rank B, will be square (of dimensions $B \times B$) and of rank B by Sylvester's inequality [13]. In this case, the resulting coefficient matrix will be invertible, and a solution is possible.

The procedure just described, although involving more computation than, for example, nodal analysis, is general. If one solves for all element currents and voltages for a general circuit, however, one must anticipate additional complexity. Furthermore, the method outlined is algorithmic and can be computer automated. The element constraint matrices A_0 and B_0 consist of stylized submatrices corresponding to each type of element. These are referred to as **stamps** in the modified nodal technique described elsewhere in this volume, and the preceding method is quite similar. The major difference is that is uses node voltage as a basis for the branch voltage space of a circuit instead of the tree voltages described previously.

References

[1] P. W. Bridgman, *The Logic of Modern Physics,* New York: Macmillan, 1927.

[2] W.-K. Chen, *Linear Networks and Systems,* Monterey, CA: Brooks-Cole, 1983.

[3] J. Choma, *Electrical Networks: Theory and Analysis,* New York: Wiley, 1985.

[4] L. P. Huelsman, *Basic Circuit Theory,* Englewood Cliffs, NJ: Prentice Hall, 1927.

[5] A. M. Davis, "A unified theory of lumped circuits and differential system based on Heaviside operators and causality," *IEEE Trans. Circuits Syst.,* vol. 41, no. 11, pp. 712–727, November, 1990.

[6] A. M. Davis, *Linear Circuit Analysis*, text in preparation.

[7] W.-K. Chen, *Applied Graph Theory: Graphs and Electrical Networks,* New York: North-Holland, 1976.

[8] Chan, Shu-Park, *Introductory Topological Analysis of Electrical Networks,* New York: Holt, Rinehart, & Winston, 1969.

[9] A. M. Davis, unpublished notes.

[10] H. J. Carlin and D. C. Youla, "Network synthesis with negative resistors," *Proc. IEEE,* vol. 49, pp. 907–920, May 1961.

[11] V. Belevitch, "Four dimensional transformations of 4-pole matrices with applications to the synthesis of reactance 4-poles," *IRE Trans. Circuit Theory,* vol. CT-3, pp. 105–111, June 1956.

[12] B. D. H. Tellegen, "La Recherche pour une Serie Complete d'Elements de Circuit Ideaux Non-Lineaires," *Rendiconti Del Seminario Mathematico e Fisico di Milano,* vol. 25, pp. 134–144, April 1954.

[13] F. R. Gantmacher, *Theory of Matrices*, New York: Chelsea, 1959.

2.2 Network Theorems

Marwan A. Simaan

In Section 2.1, we learned how to determine currents and voltages in a resistive circuit. Methods have been developed, which are based on applying Kirchhoff's voltage law (KVL) and current law (KCL), to derive a set of mesh or node equations which, when solved, will yield mesh currents or node voltages, respectively. Frequently, and especially if the circuit is complex with many elements, the application of these methods may be considerably simplified if the circuit itself is simplified. For example, we may wish to replace a portion of the circuit consisting of resistors and sources by an equivalent circuit that has fewer elements in order to write fewer mesh or node equations.

In this context, we introduce three important and related theorems known as the **superposition**, the **Thévenin** and the **Norton theorems**. The superposition theorem shows how to solve for a variable in a circuit that has many independent sources, by solving simpler circuits, each excited by only one source. The Thévenin and Norton theorems can be used to replace a portion of a circuit at any two terminals by an equivalent circuit which consists of a voltage source in series with a resistor (i.e., a nonideal voltage source) or a current source in parallel with a resistor (i.e., a nonideal current source). Another important result derived in this section concerns the calculation of power dissipated in a load resistor connected to a circuit. This result is known as the **maximum power transfer theorem**, and is frequently used in circuit design problems. Finally, a result known as the **reciprocity theorem** is also discussed.

An important property of linear resistive circuits is the type of relationship that exists between any variable in the circuit and the independent sources. For linear resistive circuits the solution for a voltage or current variable can always be expressed as a *linear combination of the independent sources*. Let us elaborate on what we mean by this statement through an example.

Consider the circuit in Figure 2.60 and assume that we are interested in the voltage v across R_2. We can solve for v by first applying KCL at node a to get the equation

$$-\frac{v-v_1}{R_1}+\beta i_x+i_1-\frac{v}{R_2}=0 \tag{2.110}$$

and then by making use of the fact that

$$i_x=\frac{v_1-v}{R_1} \tag{2.111}$$

This gives

$$v=\frac{(1+\beta)R_2}{R_1+(1+\beta)R_2}v_1+\frac{R_1R_2}{R_1+(1+\beta)R_2}i_1 \tag{2.112}$$

Here, the voltage v is a linear combination of the independent sources v_1 and i_1.

The preceding observation indeed applies to every **linear circuit**. In general, if we let y denote a voltage across, or a current in, any element in a linear circuit and if we let $\{x_1, x_2, \ldots, x_N\}$ denote the independent voltage and current sources in that circuit (assuming there is a total of N such sources), then we can write

$$y=\sum_{k=1}^{N}a_kx_k \tag{2.113}$$

where $a_1, a_2, \ldots, a_N$ are constants which depend on the circuit parameters. Thus, in the circuit of Figure 2.60, every current or voltage variable can be expressed as a linear combination of the form

$$y = a_1V_1 + a_2 i_1$$

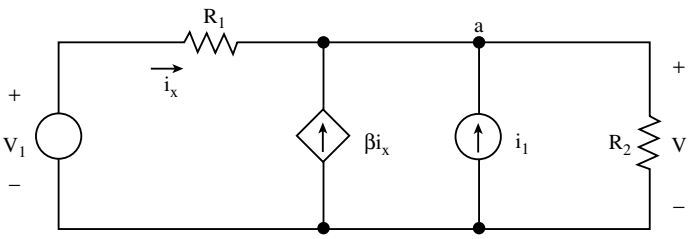


FIGURE 2.60 An example of a linear circuit.

where a_1 and a_2 are constants that depend on R_1, R_2, and β. For the voltage v across R_2, this relationship is given by expression (2.112).

Mathematically, the relationship between y and $\{x_1, x_2, \ldots, x_N\}$ expressed in (2.113) is said to be *linear* because it satisfies the following two conditions:

1. The *superposition condition,* which requires that:

$$\text{if} \qquad \hat{y} = \sum_{k=1}^{N} a_k \hat{x}_k$$

$$\text{and} \qquad \tilde{y} = \sum_{k=1}^{N} a_k \tilde{x}_k$$

$$\text{then} \qquad \hat{y} + \tilde{y} = \sum_{k=1}^{N} a_k \left(\hat{x}_k + \tilde{x}_k\right)$$

2. The *homogeneity condition,* which requires that:

$$\text{if} \qquad \hat{y} = \sum_{k=1}^{N} a_k \hat{x}_k$$

$$\text{then} \qquad c\hat{y}(t) = \sum_{k=1}^{N} a_k \{c\hat{x}_k\}$$

for any constant c.

The following example illustrates how these two conditions are satisfied for the circuit of Figure 2.60.

Example 2.1. For the circuit of Figure 2.60, let $R_1 = 2\ \Omega$, $R_2 = 1\ \Omega$, and $\beta = 2$. Show that the expression for v in terms of v_1 and i_1 satisfies the superposition and homogeneity conditions.

Substituting the values of R_1, R_2, and β in (2.112), the expression for v becomes

$$v = \frac{3}{5} v_1 + \frac{2}{5} i_1 \tag{2.114}$$

To check the superposition property, let $v_1 = \hat{v}_1$ and $i_1 = \hat{i}_1$. Then, the voltage $\hat{v}$ across R_2 is

$$\hat{v} = \frac{3}{5} \hat{v}_1 + \frac{2}{5} \hat{i}_1$$

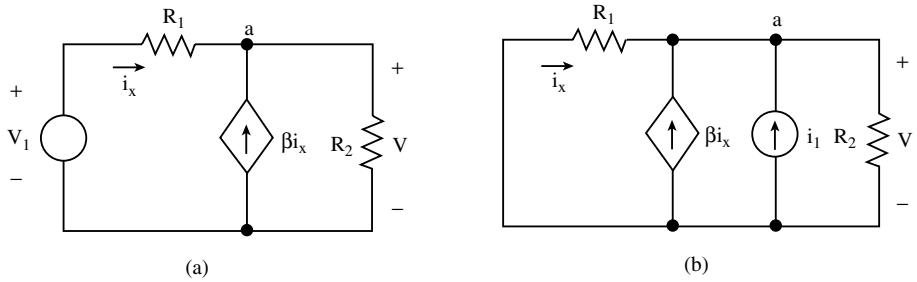


FIGURE 2.61 Circuit of Figure 2.60 with (a) the current source deactivated and with (b) the voltage source deactivated.

Similarly, let $V_1 = \tilde{V}_1$ and $i_1 = \tilde{i}_1$. Then, the voltage $\tilde{V}$ across R_2 is

$$\tilde{v} = \frac{3}{5}\tilde{v}_1 + \frac{2}{5}\tilde{i}_1$$

Now, assume that $V_1 = \hat{V}_1 + \tilde{V}_1$ and that $i_1 = \hat{i}_1 + \tilde{i}_1$. Then, according to (2.114) the corresponding voltage V across R_2 is

$$\begin{aligned} V &= \frac{3}{5}\left(\hat{V}_1 + \tilde{V}_1\right) + \frac{2}{5}\left(\hat{i}_1 + \tilde{i}_1\right) \\ &= \left(\frac{3}{5}\hat{V}_1 + \frac{2}{5}\hat{i}_1\right) + \left(\frac{3}{5}\tilde{V}_1 + \frac{2}{5}\tilde{i}_1\right) \\ &= \hat{V} + \tilde{V} \end{aligned}$$

Hence, the superposition condition is satisfied.

To check the homogeneity condition, let $V_1 = c\hat{V}_1$ and $i_1 = c\hat{i}_1$, where c is an arbitrary constant. Then, according to (2.114) the corresponding voltage v across R_2 is

$$\begin{aligned} V &= \frac{3}{5}\left(c\hat{V}_1\right) + \frac{2}{5}\left(c\hat{i}_1\right) \\ &= c\left(\frac{3}{5}\hat{V}_1 + \frac{3}{5}\hat{i}_1\right) \\ &= c\hat{V} \end{aligned}$$

The homogeneity condition is also satisfied.

The Superposition Theorem

Let us reexamine expression (2.112) for the voltage v in the circuit of Figure 2.60. To be more specific, let us use this expression to calculate the voltage across R_2 for the two circuits shown in Figures 2.61(a) and 2.61(b), respectively. Observe that the first circuit is obtained from the original circuit *by deactivating* the current source (i.e., setting $i_1 = 0$) and leaving the voltage source to act alone. The second is obtained by *deactivating* the voltage source (i.e., setting $V_1 = 0$) and leaving the current source to act alone. If we label the voltages across R_2 in these two circuits as v_a and v_b, respectively, then

$$v_a = v\Big|_{\substack{\text{when} \\ i_1=0}} = \frac{(1+\beta)R_2}{R_1 + (1+\beta)R_2}v_1$$

and

$$v_b = v\Big|_{\substack{\text{when}\\ v_1=0}} = \frac{R_1R_2}{R_1+(1+\beta)R_2}i_1$$

In other words, expression (2.112), which was used to derive the above two expressions, can itself be written as:

$$v = v\Big|_{\substack{\text{when}\\ i_1=0}} + v\Big|_{\substack{\text{when}\\ v_1=0}}$$

or

$$v = v_a + v_b$$

Thus, we conclude that the voltage across R_2 in Figure 2.60 is actually equal to the sum of two voltages across R_2 due to two independent sources in the circuit acting individually.

The preceding result is in fact a direct consequence of the linearity property

$$y = a_1x_1 + a_2x_2 + \ldots + a_Nx_N$$

expressed in (2.113). Note that, from this expression, we can write

$$a_1x_1 = y\Big|_{\text{when } x_1\neq 0, x_2=0, x_3=0,\ldots,x_N=0},$$

$$a_2x_2 = y\Big|_{\text{when } x_1=0, x_2\neq 0, x_3=0,\ldots,x_N=0},$$

$$\vdots \quad \vdots$$

$$a_Nx_N = y\Big|_{\text{when } x_1=0, x_2=0, x_3=0,\ldots,x_N\neq 0}$$

This means (2.113) can be rewritten as

$$y = y\Big|_{\text{when } x_1\neq 0, x_2=0, x_3=0,\ldots,x_N=0} + y\Big|_{\text{when } x_1=0, x_2\neq 0, x_3=0,\ldots,x_N=0} + \vdots + y\Big|_{\text{when } x_1=0, x_2=0, x_3=0,\ldots,x_N\neq 0}$$

The following theorem, known as the **superposition theorem**, is therefore directly implied from the previous expression:

> *The voltage across any element (or current through any element) in a linear circuit may be calculated by adding algebraically the individual voltages across that element (or currents through that element) due to each independent source acting alone with all other independent sources deactivated.*

In this statement, the word *deactivated* is used to imply that the source is set to zero. In this context, we refer to a **deactivated current source** as one that is replaced by an open circuit and a **deactivated**

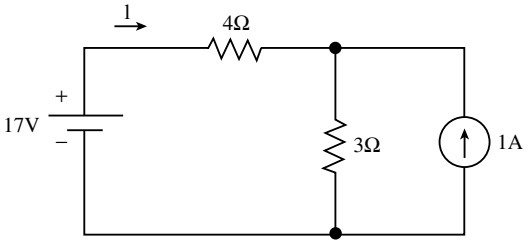


FIGURE 2.62 Circuit for Example 2.2.

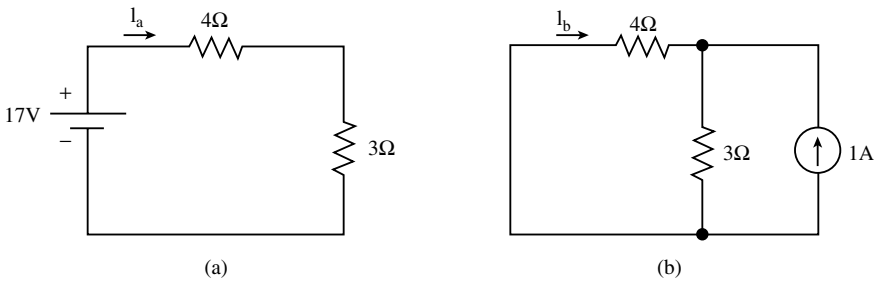


FIGURE 2.63 Circuit for Example 2.2 with (a) the current source deactivated and with (b) the voltage source deactivated.

voltage source as one that is replaced by a short circuit. Note that the action of deactivating a source refers only to independent sources. The following example illustrates how the superposition theorem can be used to solve for a variable in a circuit with more than one independent source.

Example 2.2. For the circuit shown in Figure 2.62, apply superposition to calculate the current I in the 4 Ω resistor.

Because we are interested in calculating I using the superposition theorem, we need to consider the two circuits shown in Figures 2.63(a) and (b). The first is obtained by deactivating the current source and the second is obtained by deactivating the voltage source. Let I_a be the current in the 4 Ω resistor in the first circuit and I_b be the current in the same resistor in the second. Then, by superposition

$$I = I_a + I_b$$

We can solve for I_a and I_b independently as follows. From Figure 2.63(a):

$$I_a = \frac{17}{7} A$$

and from Figure 2.63(b), applying the current divider rule,

$$I_b = -1 \cdot \frac{3}{7} A$$

Thus,

$$I = \frac{17}{7} - \frac{3}{7}$$

$$= 2A$$

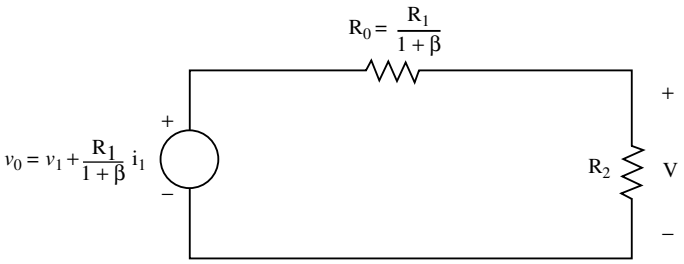


FIGURE 2.64 Voltage divider circuit representing (2.118).

The Thévenin Theorem

In the discussion on the superposition theorem, we interpreted (2.112) for the voltage V in the circuit of Figure 2.60 as a superposition of two terms. Let us now examine a different interpretation of this expression.

Suppose we factor the common term in expression (2.112), so that it can be written as

$$v = \frac{(1+\beta)R_2}{R_1+(1+\beta)R_2}\left\{v_1 + \frac{R_1}{1+\beta} i_1\right\} \tag{2.115a}$$

or

$$v = \frac{R_2}{\left(R_1/(1+\beta)\right)+R_2}\left\{v_1 + \left(\frac{R_1}{1+\beta}\right) i_1\right\} \tag{2.115b}$$

Now, suppose we define

$$v_0 = v_1 + \frac{R_1}{1+\beta} i_1 \tag{2.116}$$

and

$$R_0 = \frac{R_1}{1+\beta} \tag{2.117}$$

Then, we can write (2.112) in the simple form

$$v = \frac{R_2}{R_0+R_2} v_0 \tag{2.118}$$

This expression can be interpreted as a voltage divider equation for a two-resistor circuit as shown in Figure 2.64. This circuit has a voltage source v_0 in series with two resistors R_0 and R_2. When this circuit is compared with Figure 2.60, the combination of voltage source v_0 in series with R_0 can be interpreted as an equivalent replacement of all the elements in the circuit connected to R_2. That is, we could remove that portion of the circuit of Figure 2.60 consisting of v_1, R_1, βi_x, and i_1 and replace it with the voltage source v_0 in series with the resistor R_0.

The fact that a portion of a circuit can be replaced by an equivalent circuit consisting of a voltage source in series with a resistor is actually a direct result of the linearity property, and hence is true for linear circuits in general. It is known as **Thévenin's theorem**[11] and is stated as follows:

[11]For an interesting, brief discussion on the history of Thévenin's theorem, see an article by James E. Brittain, in *IEEE Spectrum,* p. 42, March 1990.

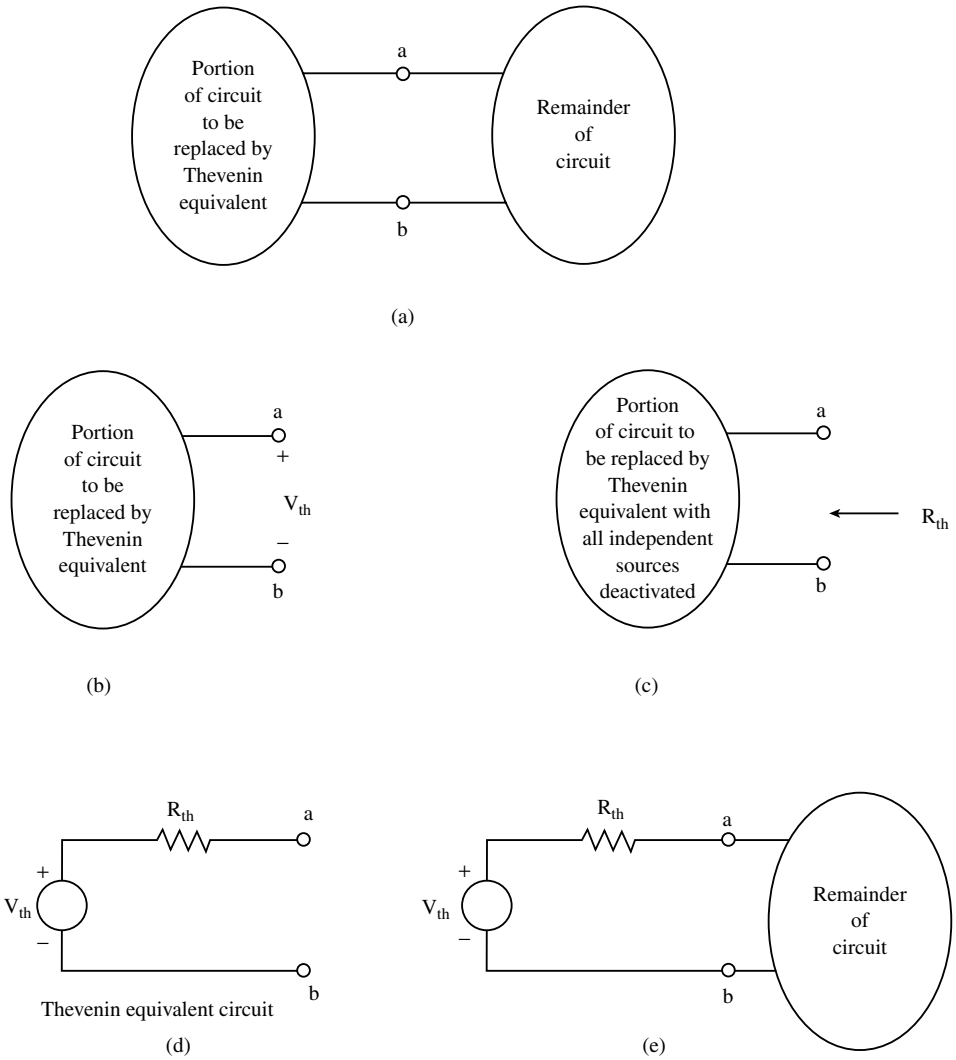


FIGURE 2.65 Steps in determining the Thévenin equivalent circuit.

Any portion of a linear circuit between two terminals a and b can be replaced by an equivalent circuit consisting of a voltage source V_{th} in series with a resistor R_{th}. The voltage V_{th} is determined as the open circuit voltage at terminals a-b. The resistor R_{th} is equal to the input resistance at terminals a-b with all the independent sources deactivated.

The various steps involved in the derivation of the Thévenin equivalent circuit are illustrated in Figure 2.65. The Thévenin voltage v_{th} is determined by solving for the voltage at terminals a-b when open circuited, and the Thévenin resistance R_{th} is determined by calculating the input resistance of the circuit at terminals a-b when all the independent sources have been deactivated. The following two examples illustrate the application of this important theorem.

Example 2.3. For the circuit in Figure 2.66, determine the Thévenin equivalent of the portion of the circuit to the left of terminals a-b; use it to calculate the current I in the 2 Ω resistor.

First, we determine V_{th} from the circuit of Figure 2.67(a). Note that because terminals a-b are open circuited the current in the branch containing the 2 Ω resistor and 9 V source is equal to 3 A in the direction shown. Applying KCL at the upper node of the 1 Ω resistor, we can calculate the current in

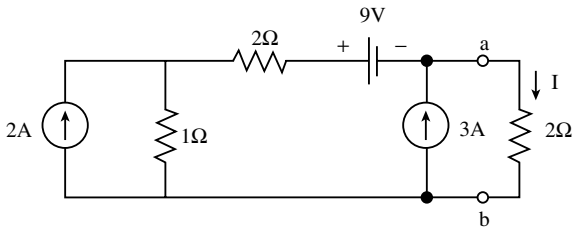


FIGURE 2.66 Circuit for Example 2.3.

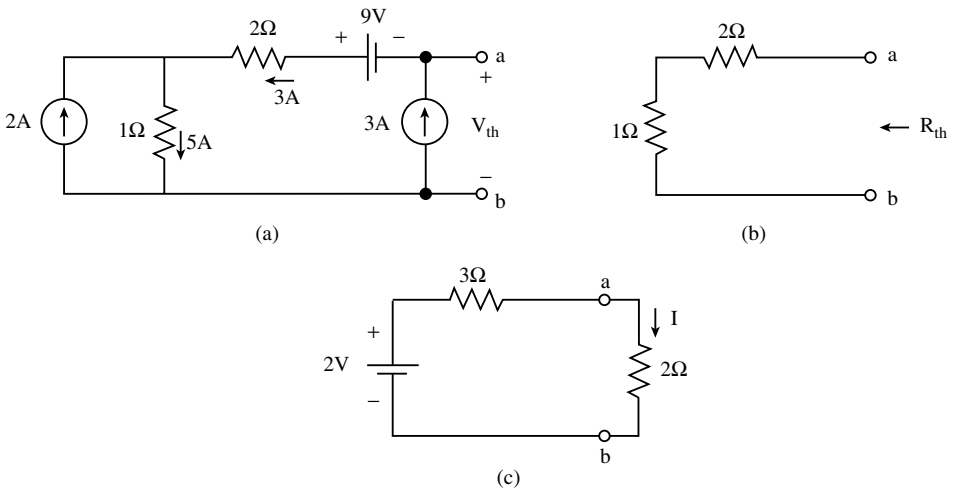


FIGURE 2.67 (a) Calculation of V_{th}, (b) calculation of R_{th}, and (c) the equivalent circuit for Example 2.3.

this resistor to be 3 + 2 = 5 A as shown. Writing a KVL equation around the inner loop (counterclockwise at terminal b) we have

$$V_{th} + 9 - (2 \times 3) - (1 \times 5) = 0$$

which yields

$$V_{th} = 2 \text{ V}$$

Now, for R_{th} the three sources are deactivated to obtain the circuit shown in Figure 2.67(b). From this circuit, it is clear that

$$R_{th} = 3\Omega$$

The circuit obtained by replacing the portion to the left of terminals a-b with its Thévenin equivalent is shown in Figure 2.67(c). The current I is now easily computed as

$$I = \frac{2}{2+3} = 0.4A$$

Example 2.4. For the circuit shown in Figure 2.60, determine the Thévenin equivalent circuit for the portion of the circuit to the left of resistor R_2.

In deriving (2.118) from (2.110), we actually already determined the Thévenin equivalent for the portion of the circuit to the left of R_2. This was shown in Figure 2.64. Of course, this procedure is *not* the most efficient way to determine the Thévenin equivalent. Let us now illustrate how the equivalent

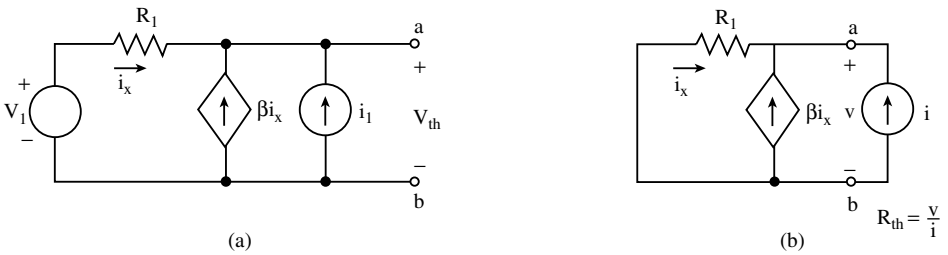


FIGURE 2.68 Calculation of (a) v_{th} and (b) R_{th} for the circuit of Figure 2.60 (Example 2.4).

circuit is obtained using the procedure described in Thévenin's theorem. First, we determine v_{th} from the circuit of Figure 2.68(a) with R_2 removed and terminals a-b left open. Applying KCL at node a, we have

$$i_x + \beta i_x + i_1 = 0$$

or

$$i_x = -\frac{i_1}{1+\beta}$$

Hence,

$$v_{th} = v_1 - R_1 i_x$$
$$= v_1 + \frac{R_1}{1+\beta} i_1$$

As for R_{th}, we need to consider the circuit shown in Figure 2.68(b), in which the two independent sources were deactivated. Because of the presence of the dependent source βi_x, we determine R_{th} by exciting the circuit with an external source. Let us use a current source i for this purpose and determine the voltage v across it as shown in Figure 2.68(b). We stress that i is an arbitrary and completely independent source and is in no way related to i_1, which was deactivated. Applying KCL at node a, we have

$$i_x + \beta i_x + i = 0$$

or

$$i_x = -\frac{1}{1+\beta} i$$

Also, applying Ohm's law to R_1,

$$v = -R_1 i_x$$

or

$$v = \frac{R_1}{1+\beta} i$$

Hence,

$$R_{th} = \frac{v}{i}$$
$$= \frac{R_1}{1+\beta}$$

Note that v_{th} and R_{th} determined previously are the same as v_0 and R_0 of (2.116) and (2.117).

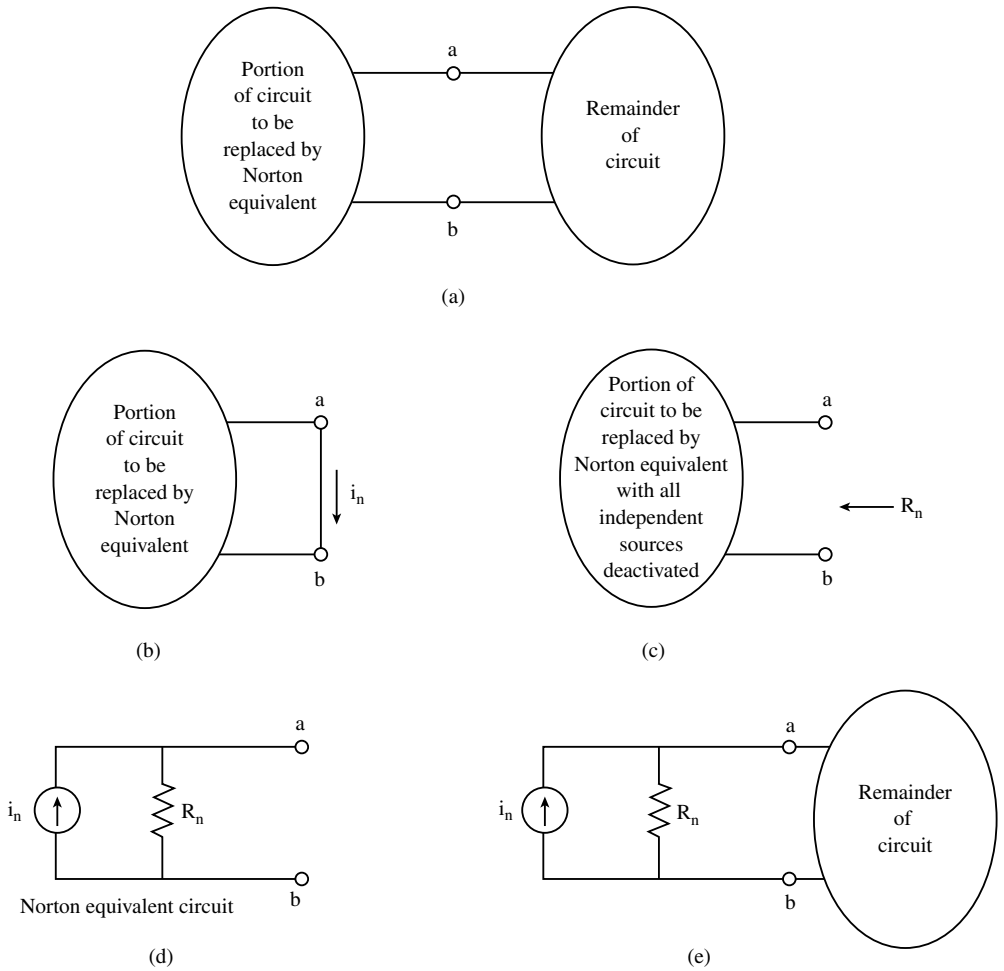


FIGURE 2.69 Steps in determining the Norton equivalent circuit.

The Norton Theorem

Instead of a voltage source in series with a resistor, it is possible to replace a portion of a circuit by an equivalent current source in parallel with a resistor. This result is formally known as *Norton's theorem* and is stated as follows:

> *Any portion of linear circuit between two terminals a and b can be replaced by an equivalent circuit consisting of a current source i_n in parallel with a resistor R_n. The current i_n is determined as the current that flows from a to b in a short circuit at terminals a-b. The resistor R_n is equal to the input resistance at terminals a-b with all the independent sources deactivated.*

As in the case of Thévenin's, the preceding theorem provides a procedure for determining the "Norton" current source and "Norton" resistance in the **Norton equivalent circuit**. The various steps in this procedure are illustrated in Figure 2.69. The Norton current is determined by solving for the current in a short circuit at terminals a-b and the Norton resistance is determined by calculating the input resistance to the circuit at terminals a-b when all independent sources have been deactivated. The Norton's equivalent of a portion of a circuit is, in effect, a nonideal current source representation of that portion. It should be noted that the procedure to determine R_n is exactly the same as that for R_{th}. In other words,

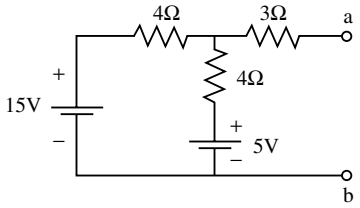


FIGURE 2.70 Circuit for Example 2.5.

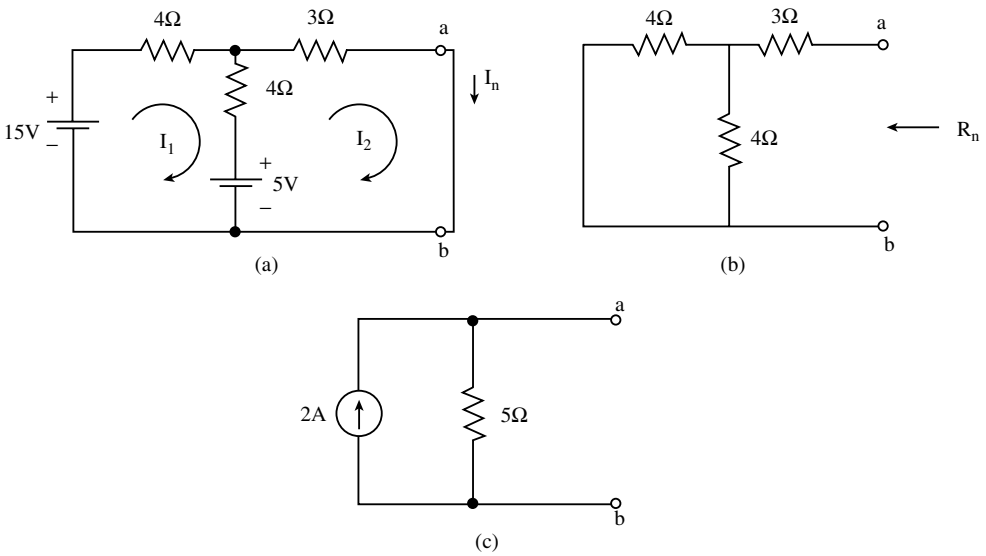


FIGURE 2.71 (a) Calculation of I_n, (b) calculation of R_n and (c) Norton's equivalent for the circuit of Example 2.5.

$$R_n = R_{th}$$

Also, if we compare Thévenin's and Norton's equivalent circuits, we see that these are indeed related by the voltage–current source transformation rule discussed earlier in this chapter. Each circuit is a source transformation of the other. For this reason, the Thévenin and Norton equivalent circuits are often referred to as **dual circuits**, and the two resistance R_{th} and R_n are frequently referred to as the source resistance and denoted by R_s. Clearly, v_{th} and i_n are related by

$$v_{th} = R_s i_n$$

Example 2.5. For the circuit shown in Figure 2.70, determine the Norton equivalent circuit at terminals a-b.

We determine Norton's current I_n by placing a short circuit between a and b, as shown in Figure 2.71(a), and solving for the current in it with a reference direction going from a to b. For this circuit, we could use the mesh equation method. In matrix form, the mesh equations are

$$\begin{bmatrix} 8 & -4 \\ -4 & 7 \end{bmatrix} \cdot \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \begin{bmatrix} 10 \\ 5 \end{bmatrix}$$

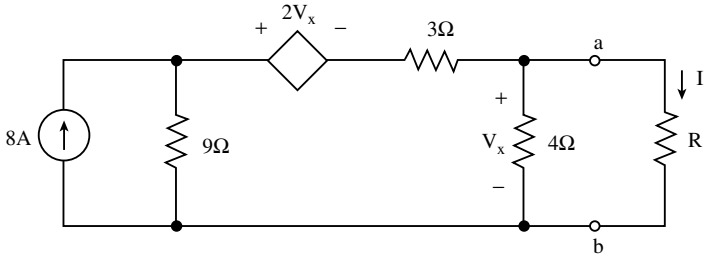


FIGURE 2.72 Circuit for Example 2.6.

and the solution for the mesh currents is

$$\begin{bmatrix} I_1 \\ I_2 \end{bmatrix} = \frac{1}{56-16}\begin{bmatrix} 7 & 4 \\ 4 & 8 \end{bmatrix}\cdot\begin{bmatrix} 10 \\ 5 \end{bmatrix}$$

From this, we extract I_n as

$$I_n = I_2 = \frac{40+40}{40} = 2A$$

Norton's resistance R_n is determined by deactivating the two voltage sources, as shown in Figure 2.71(b), and calculating the input resistance at terminals a-b. Clearly,

$$R_n = \left(4\|4\right)+3 = 5\Omega$$

Thus, the Norton equivalent for the circuit of Figure 2.70 is shown in Figure 2.71(c).

Example 2.6. For the circuit shown in Figure 2.72 determine the Norton equivalent circuit at terminals a-b and use it to calculate the current and power dissipated in *R*.

With a short circuit placed at terminals a-b, as shown in Figure 2.73(a), the voltage $V_x = 0$. Hence, the dependent source in this circuit is equal to zero. This means that the 8 A current source has the 9 Ω and 3 Ω resistors in parallel across it, and I_n is the current in the 3 Ω resistor. Using the current divider rule we have

$$I_n = 8\frac{9}{12} = 6A$$

Now, deactivating the independent source to determine R_n,we excite the circuit with a voltage source *V* at terminals a-b. Let *I* be the current in this source as shown in Figure 2.73(b). Applying KCL at node a yields the current in the 3 Ω resistor to be I – (V/4) from right to left. Applying KVL around the outer loop and making use of the fact that in this circuit $V_x = V$, we get

$$V - 3\left(I - \frac{V}{4}\right) + 2V - 9\left(I - \frac{V}{4}\right) = 0$$

Solution of this equation yields

$$R_n = \frac{V}{I} = 2\Omega.$$

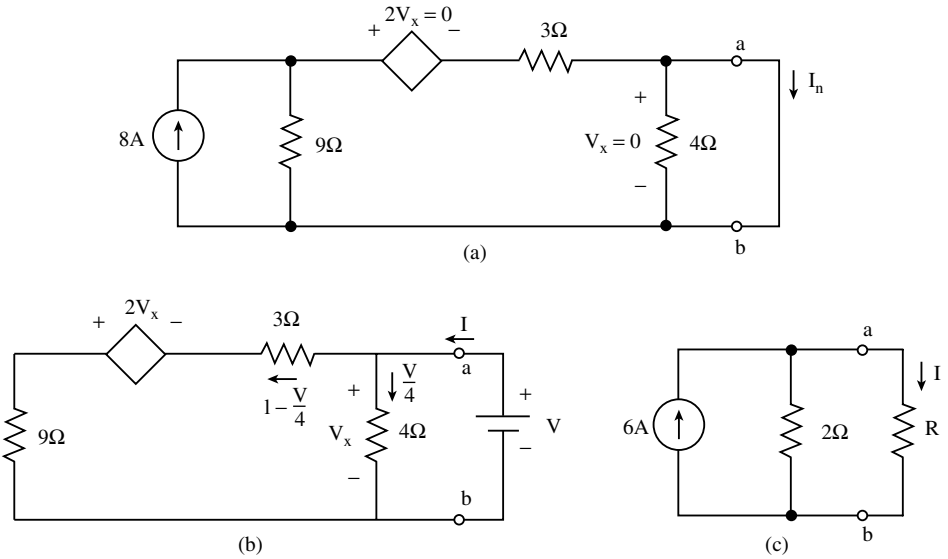


FIGURE 2.73 (a) Calculation of I_n, (b) calculation of R_n, and (c) Norton's equivalent for the circuit of Example 2.6.

The Norton equivalent of the portion of the circuit to the left of terminals a–b, connected to the resistor R is shown in Figure 2.73(c). Applying the current divider rule gives

$$I = 6\frac{2}{2+R}$$

$$= \frac{12}{2+R}\text{A}$$

and the power dissipated in R is

$$P = RI^2$$

$$P = \frac{144R}{(2+R)^2}\text{W} \tag{2.119}$$

The Maximum Power Transfer Theorem

In the previous example, we replaced the entire circuit connected to the resistor R at terminals a-b by its Norton equivalent in order to calculate the power P dissipated in R. Because R did not have a fixed value, we determined an expression for P in terms of R. Suppose we are now interested in examining how P varies as a function of R. A plot of P versus R as given by (2.119) is given in Figure 2.74.

The first noticeable characteristic of this plot is that it has a *maximum*. Naturally, we would be interested in the value of R that results in maximum power delivered to it. This information is directly available from the plot in Figure 2.74. To maximize P the value of R should be 2 **Ω** and the maximum power is P_{max} = 18 W. That is, 18 W is the most that this circuit can deliver at terminals a-b, and that occurs when R = 2 **Ω**. Any other value of R will result in less power delivered to it.

The problem of finding the value of a load resistor R_L such that maximum power is delivered to it is obviously an important circuit design problem. Because it is possible to reduce any linear circuit connected to R_L into either its Thévenin or Norton equivalent, as illustrated in Figure 2.75, the problem becomes quite simple. We need to consider only either the circuit of Figure 2.75(b) or that of 2.75(c).

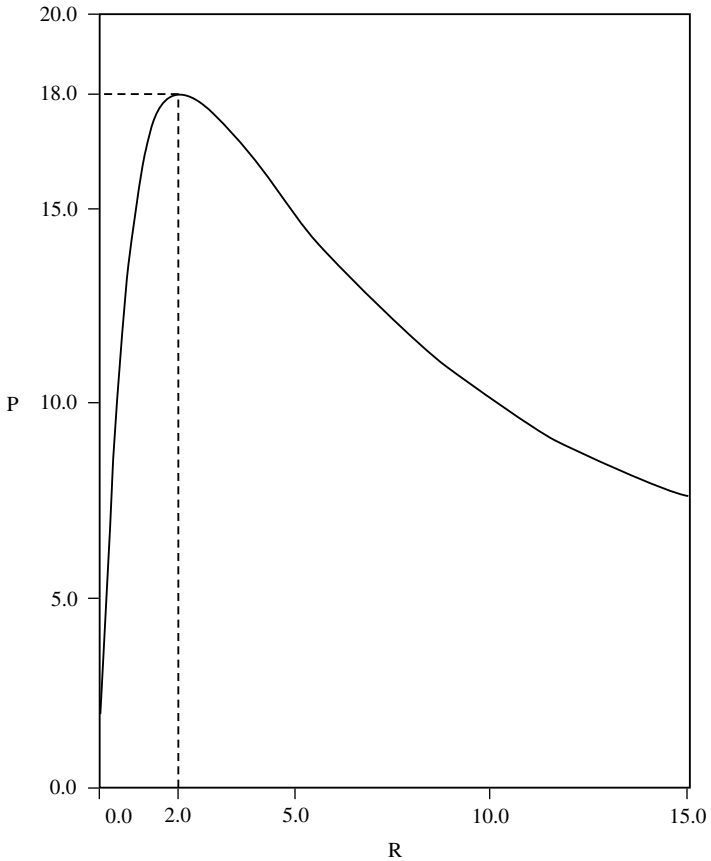


FIGURE 2.74 Plot of P vs. R for the circuit of Example 2.6.

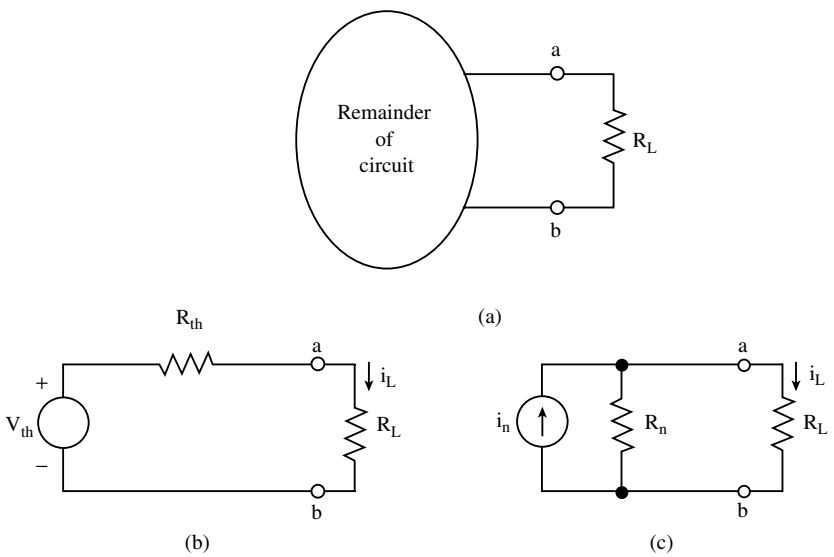


FIGURE 2.75 (a) A load resistance R_L in a circuit. (b) R_L with the remainder of the circuit reduced to a Thévenin equivalent. (c) R_L with the remainder of the circuit reduced to a Norton equivalent.

Let us first consider the circuit that uses the Thévenin equivalent. In this case, the power P delivered to R_L is given by

$$P = \left(\frac{v_{th}}{R_{th} + R_L} \right)^2 R_L \tag{2.120}$$

In general, we may not always be able to plot P vs. R_L, as we did earlier, therefore, we need to maximize P mathematically. We do this by solving the necessary condition

$$\frac{dP}{dR_L} = 0 \tag{2.121}$$

for R_L. To guarantee that R_L maximizes P, it must also satisfy the sufficiency condition

$$\left. \frac{d^2P}{dR_L^2} \right|_{R_L} < 0 \tag{2.122}$$

Thus, applying these conditions to (2.120), we have

$$\begin{aligned} \frac{dP}{dR_L} &= v_{th}^2 \left[\frac{\left(R_h + R_L\right)^2 - 2R_L\left(R_{th} + R_L\right)}{\left(R_{th} + R_L\right)^4} \right], \\ &= v_{th}^2 \frac{\left(R_{th} - R_L\right)}{\left(R_{th} + R_L\right)^3} \end{aligned} \tag{2.123}$$

Equating the right-hand side of (2.123) to zero and solving for R_L yields

$$R_L = R_{th}$$

The sufficiency condition (2.122) yields

$$\frac{d^2P}{dR_L^2} = v_{th}^2 \frac{2R_{th} - 4R_L}{\left(R_{th} + R_L\right)^4}$$

When R_L is replaced with R_{th}, we get

$$\left. \frac{d^2P}{dR_L^2} \right|_{R_L = R_{th}} = -\frac{v_{th}^2}{8R_{th}^3} < 0$$

Thus, $R_L = R_{th}$ satisfies both conditions (2.121) and (2.122), and hence is the maximizing value. This result is often referred to as the **maximum power transfer theorem**. It says

The maximum power that can be transferred to a load resistance R_L by a circuit represented by its Thévenin equivalent is attained when R_L is equal to R_{th}.

The corresponding value of $P_{\max}$ is obtained from (2.120) as

$$P_{\max} = \frac{v_{th}^2}{4R_{th}} \tag{2.124}$$

In the case of Norton's equivalent circuit of Figure 2.75(b), a similar derivation can be carried out. The power P delivered to R_L is given by

$$P = \left(\frac{R_n i_n}{R_n + R_L} \right)^2 R_L \tag{2.125}$$

This expression has exactly the same form as (2.120). Consequently, its maximum is achieved when

$$R_L = R_n$$

and the corresponding maximum power is

$$P_{\max} = \frac{R_n i_n^2}{4} \tag{2.126}$$

This leads to the following alternate statement of the **maximum power transfer theorem**:

> *The maximum power that can be transferred to a load resistance R_L by a circuit represented by its Norton equivalent is attained when R_L is equal to R_n.*

Example 2.7. Consider the circuit of Example 2.3 in Figure 2.66. Determine the value of a load resistor R_L connected in place of the 2 Ω resistor at terminals a-b in order to achieve maximum power transfer to the load.

Solution. The Thévenin equivalent for the circuit of Figure 2.66 already was determined and is shown in Figure 2.67 (c). Using the results of the maximum power transfer theorem, we should have

$$R_L = 3\ \Omega$$

The corresponding value of maximum power is

$$P_{\max} = \frac{2^2}{4 \times 3}$$

$$= \frac{1}{3} \text{W}$$

The Reciprocity Theorem

The reciprocity theorem is an important result that applies to circuits consisting of linear resistors and one independent source (either a current source or a voltage source). It does not apply to nonlinear circuits, and, in general, it does not apply to circuits containing dependent sources. The reciprocity theorem is stated as follows:

> *The ratio of a voltage (or current) response in one part of the circuit to the current (or voltage)* source is the same if the locations of the response and the source are interchanged.

It is important to note that the reciprocity theorem applies only to circuits in which the source and response are voltage and current or current and voltage, respectively. It does not apply to circuits in which the source and response are of the same type (i.e., voltage and voltage or current and current).

Example 2.8. Verify the reciprocity theorem for the circuit shown in Figure 2.76(a).

Solution. If we interchange the location of the 40 V voltage source and current response I, we obtain the circuit shown in Figure 2.76(b). The reciprocity theorem implies that I should be the same in both circuits.

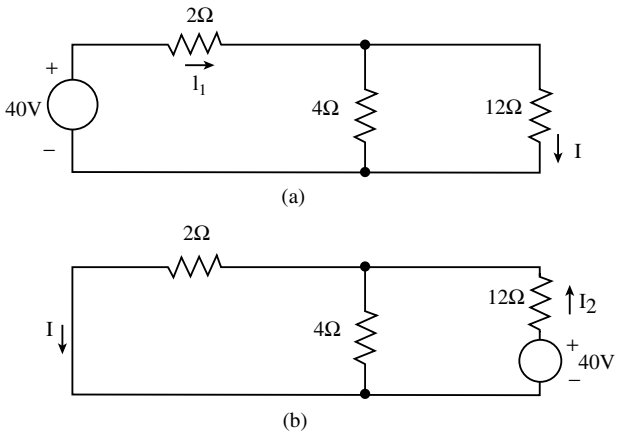


FIGURE 2.76 (a) Circuit for Example 2.8. (b) Circuit with location of voltage source and current response interchanged.

For the circuit in Figure 2.76(a), the current I_1 in the 2 Ω resistor is equal to:

$$I_1 = \frac{40}{2 + 4\|12}$$

$$= \frac{40}{2+3}$$

$$= 8\text{A}$$

and the response I can be easily determined, using the current divider rule, as:

$$I = 8 \times \frac{4}{4+12}$$

$$= 2\text{A}$$

For the circuit in Figure 2.76(b), the current I_2 in the 12 Ω resistor is equal to:

$$I_2 = \frac{40}{12 + 2\|4}$$

$$= \frac{40}{12 + \frac{4}{3}}$$

$$= 3\text{ A}$$

and the response I can be determined easily, using the current divider rule, as

$$I = 3 \times \frac{4}{2+4}$$

$$= 2\text{ A}$$

Thus, the reciprocity theorem is satisfied.

3

Terminal and Port Representations

James A. Svoboda
Clarkson University, New York

3.1 Introduction

Frequently, it is useful to decompose a large circuit into subcircuits and to consider the subcircuits separately. Subcircuits are connected to other subcircuits using terminals or ports. Terminal and port representations of a subcircuit describe how that subcircuit will act when connected to other subcircuits. Terminal and port representations do not provide the details of mesh or node equations because these details are not required to describe how a subcircuit interacts with other subcircuits.

In this chapter, terminal and port representations of circuits are described. Particular attention is given to the distinction between terminals and ports. Applications show the usefulness of terminal and port representations.

3.2 Terminal Representations

Figure 3.1 illustrates a subcircuit that can be connected to other subcircuits using terminals. The subcircuit is shown symbolically on the left and an example is shown on the right. Nodes *a*, *b*, *c*, and *d* are terminals and may be used to connect this circuit to other circuits. Node *e* is internal to the circuit and is not a terminal. The terminal voltages V_a, V_b, V_c, and V_d are node voltages with respect to an arbitrary reference node. The terminal currents I_a, I_b, I_c, and I_d describe currents that will exist when this subcircuit is connected to other subcircuits. Terminal representations show how the terminal voltages and currents are related. Several equivalent representations are possible, depending on which of the terminal currents and voltages are selected as independent variables.

A terminal representation of the example network in Figure 3.1 can be obtained by writing a node equation for the network. A simpler procedure is available for passive networks, i.e., networks consisting entirely of resistors, capacitors, and inductors. Because the circuit in Figure 3.1 is a passive circuit, the simpler procedure will be used to write terminal equations to represent this circuit.

The nodal admittance matrix of the example circuit will have five rows and five columns. The rows, and also the columns, of this matrix will correspond to the five nodes of the circuit in the order *a*, *b*, *c*, *d* and *e*. For example, the fourth row of the nodal admittance matrix corresponds to node *d* and the third column corresponds to node *c*. Let y_{ij} denote the admittance of a branch of the network which is incident with nodes *i* and *j* and let y_{ij} denote the element of the nodal admittance matrix that is in row *i* and column *j*. Then,

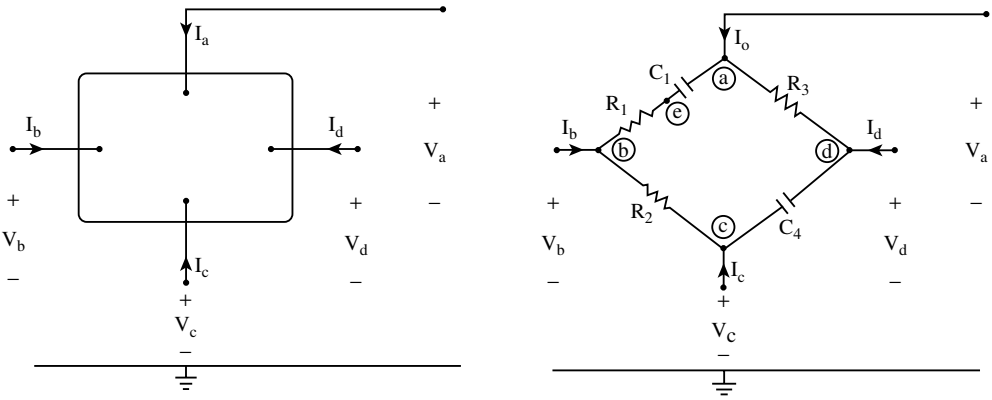


FIGURE 3.1 A four-terminal network.

$$Y_{ii} = \sum_{\substack{\text{all branches} \\ \text{incident to node } i}} y_{ik} \tag{3.1}$$

i.e., the diagonal element of the nodal admittance matrix in row i and column i is equal to the sum of the admittances of all branches in the circuit that are incident with node i. The off-diagonal elements of the nodal admittance matrix are given by

$$Y_{ij} = -\sum_{\substack{\text{all branches incident} \\ \text{to both nodes } i \text{ and } j}} y_{ij} \tag{3.2}$$

The example circuit is represented by the node equation

$$\begin{pmatrix} I_a \\ I_b \\ I_c \\ I_d \\ 0 \end{pmatrix} = \begin{pmatrix} C_1 s + \frac{1}{R_3} & 0 & 0 & -\frac{1}{R_3} & -C_1 s \\ 0 & \frac{1}{R_1} + \frac{1}{R_2} & -\frac{1}{R_2} & 0 & -\frac{1}{R_1} \\ 0 & -\frac{1}{R_2} & C_4 s + \frac{1}{R_2} & -C_4 s & 0 \\ -\frac{1}{R_3} & 0 & -C_4 s & C_4 s + \frac{1}{R_3} & 0 \\ -C_1 s & -\frac{1}{R_1} & 0 & 0 & C_1 s + \frac{1}{R_1} \end{pmatrix} \begin{pmatrix} V_a \\ V_b \\ V_c \\ V_d \\ V_e \end{pmatrix} \tag{3.3}$$

Suppose that $C_1 = 1$ F, $C_4 = 2$ F, $R_1 = 1/2\ \Omega$, $R_2 = 1/4\ \Omega$, and $R_3 = 1\ \Omega$. Then,

$$\begin{pmatrix} I_a \\ I_b \\ I_c \\ I_d \\ 0 \end{pmatrix} = \begin{pmatrix} s+1 & 0 & 0 & -1 & -1s \\ 0 & 6 & -4 & 0 & -2 \\ 0 & -4 & 2s+4 & -2s & 0 \\ -1 & 0 & -2s & 2s+1 & 0 \\ -s & -2 & 0 & 0 & s+2 \end{pmatrix} \begin{pmatrix} V_a \\ V_b \\ V_c \\ V_d \\ V_e \end{pmatrix}$$

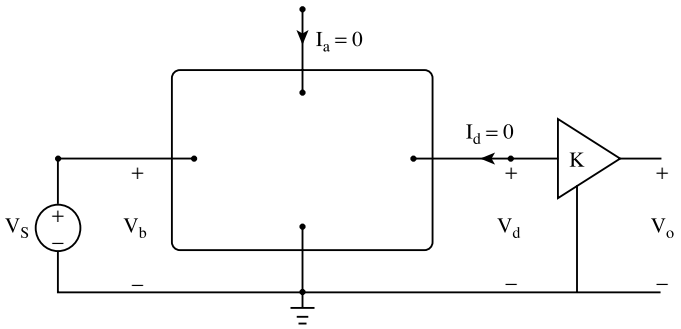


FIGURE 3.2 First application of the four-terminal network.

No external current I_e exists because node *e* is not a terminal. Row 5 of this equation, corresponding to node *e*, can be solved for V_e and then V_e can be eliminated. Doing so results in

$$\begin{pmatrix} I_a \\ I_b \\ I_c \\ I_d \end{pmatrix} = \begin{pmatrix} \frac{3s+2}{s+2} & -\frac{2s}{s+2} & 0 & -1 \\ -\frac{2s}{s+2} & \frac{6s+8}{s+2} & -4 & 0 \\ 0 & -4 & 2s+4 & -2s \\ -1 & 0 & -2s & 2s+1 \end{pmatrix} \begin{pmatrix} V_a \\ V_b \\ V_c \\ V_d \end{pmatrix} \tag{3.4}$$

The terminal voltages were selected to be the independent variables and the entries in the matrix are admittances. Because none of the nodes of the subcircuit was chosen to be the reference node, the rows and columns of the matrix both sum to zero. This matrix is called an indefinite admittance matrix. Because it is singular, Eq. (3.4) cannot be solved directly. To see the utility of Eq. (3.4), consider Figure 3.2. Here, the four-terminal network has been connected to a voltage source and an amplifier and node *c* has been grounded. This external circuitry restricts the terminal currents and voltages of the four terminal network. In particular,

$$V_b = V_s, \quad I_a = 0, \quad V_c = 0, \quad I_d = 0 \quad \text{and} \quad V_o = KV_d \tag{3.5}$$

Under these conditions, Eq. (3.4) becomes

$$\begin{pmatrix} 0 \\ I_b \\ I_c \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{3s+2}{s+2} & -\frac{2s}{s+2} & 0 & -1 \\ -\frac{2s}{s+2} & \frac{6s+8}{s+2} & -4 & 0 \\ 0 & -4 & 2s+4 & -2s \\ -1 & 0 & -2s & 2s+1 \end{pmatrix} \begin{pmatrix} V_a \\ V_s \\ 0 \\ \frac{V_o}{K} \end{pmatrix} \tag{3.6}$$

Because I_b and I_c are not of interest, the second and third rows of this equation can be ignored. The first and fourth rows can be solved to obtain the transfer function of this circuit

$$\frac{V_o(s)}{V_s(s)} = \frac{K}{3(s+1)} \tag{3.7}$$

Next, consider Figure 3.3. The four terminal network is used in a second, different application. In this case,

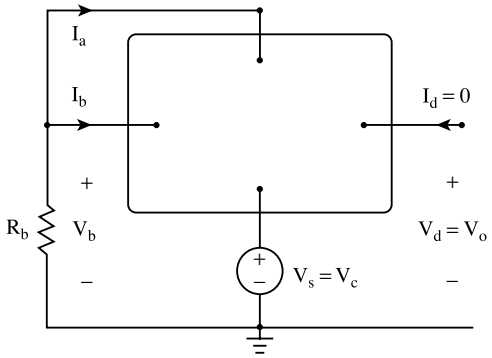


FIGURE 3.3 Second application of the four-terminal network.

$$I_a + I_b = -\frac{V_b}{R_b}, \quad V_c = V_s, \quad V_a = V_b \quad I_d = 0 \quad \text{and} \quad V_d = V_o \tag{3.8}$$

Because $V_a = V_b$, the first two columns of Eq. (3.4) can be added together, replacing V_a by V_b. Also, it is convenient to add the first two rows together to obtain $I_a + I_b$. Then, Eq. (3.4) reduces to

$$\begin{pmatrix} -\dfrac{V_b}{R_b} \\ I_c \\ 0 \end{pmatrix} = \begin{pmatrix} 5 & -4 & -1 \\ -4 & 2s+4 & -2s \\ -1 & -2s & 2s+1 \end{pmatrix} \begin{pmatrix} V_b \\ V_s \\ V_o \end{pmatrix} \tag{3.9}$$

Because I_c is not of interest, the second row of this equation can be ignored. The first and third rows of this equation can be solved to obtain the transfer function

$$\frac{V_o(s)}{V_s(s)} = \frac{(10R_b + 2)s + 4R_b}{(10R_b + 2)s + 4R_b + 1} \tag{3.10}$$

These examples illustrate the utility of the terminal equations. In these examples, the problem of analyzing a network was divided into two parts. First, the terminal equation was obtained from the node equation by eliminating the rows and columns corresponding to nodes that are not terminals. Second, the terminal equation is combined with equations describing the external circuitry connected to the four terminal network. When the external circuit is changed, only the second step must be redone. The advantage of representing a subnetwork by terminal equations is greater when

1. The subnetwork has many nodes that are not terminals.
2. The subnetwork is expected to be a component of many different networks.

3.3 Port Representations

A port consists of two terminals with the restriction that the terminal currents have the same magnitude but opposite sign. Figure 3.4 shows a two-port network. In this case, the two-port network was constructed from the four-terminal network by pairing terminals *a* and *b* form port 1 and pairing terminals *d* and *c* to form port 2. The restrictions

$$I_1 = I_a = -I_b \quad \text{and} \quad I_2 = I_d = -I_c \tag{3.11}$$

must be satisfied in order for these two pairs of terminals to be ports.

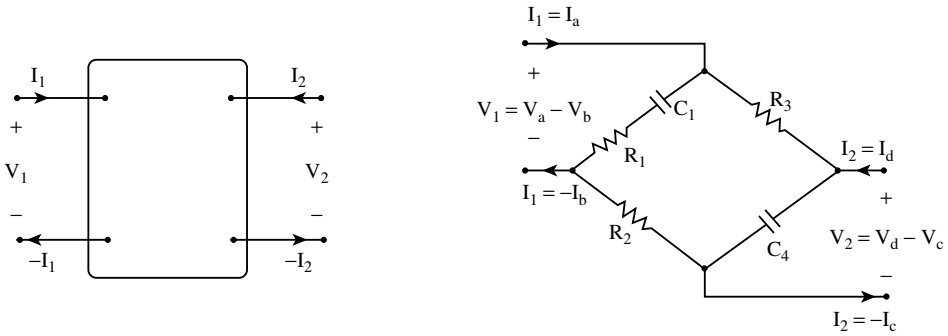


FIGURE 3.4 A two-port network.

The behavior of the two-port network is described using four variables. These variables are the port voltages, V_1 and V_2 and the port currents I_1, and I_2. Several equivalent representations are possible, depending on which two of the port currents and voltages are selected as independent variables. The left column of Table 3.1 shows the two-port representations corresponding to four of the possible choices of independent variables. These representations are equivalent, and the right column of Table 3.1 shows how one representation can be obtained from another.

In row 1 of Table 3.1, the port voltages are selected to be the independent variables. In this case, the two-port circuit is represented by an equation of the form

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \tag{3.12}$$

The elements of the matrix in this equation have units of admittance. They are denoted using the letter y to suggest admittance and are called the "y parameters" or the "admittance parameters" of the two-port network.

In row 2 of Table 3.1, the port currents are selected to be the independent variables. Now the elements of the matrix have units of impedance. They are denoted using the letter z to suggest impedance and are called the "z parameters" or the "impedance parameters" of the two-port network.

In row 3 of Table 3.1, I_1 and V_2 are the independent variables. The elements of the matrix do not have the same units. In this sense, this is a hybrid matrix. They are denoted using the letter h to suggest hybrid and are called the "h parameters" or the "hybrid parameters" of the two-port network. Hybrid parameters are frequently used to represent bipolar transistors.

In the last row of Table 3.1, the independent signals are V_2 and $-I_2$. In this case, the two-port parameters are called "transmission parameters" or "*ABCD* parameters". They are convenient when two-port networks are connected in cascade.

Next, consider the problem of calculating or measuring the parameters of a two-port circuit. Equation (3.12) suggests a procedure for calculating the y parameters of a two-port circuit. The first row of Eq. (3.12) is

$$I_1 = y_{11}V_1 + y_{12}V_2 \tag{3.13}$$

Setting $V_1 = 0$ leads to

$$y_{12} = \frac{I_1}{V_2} \quad \text{when} \quad V_1 = 0 \tag{3.14}$$

TABLE 3.1 Relationships between Two-Port Representations

$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$ $\lvert Y \rvert = y_{11}y_{22} - y_{12}y_{21}$	$\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} = \begin{pmatrix} \frac{z_{22}}{\lvert Z \rvert} & -\frac{z_{12}}{\lvert Z \rvert} \\ -\frac{z_{21}}{\lvert Z \rvert} & \frac{z_{11}}{\lvert Z \rvert} \end{pmatrix} = \begin{pmatrix} \frac{1}{h_{11}} & -\frac{h_{12}}{h_{11}} \\ \frac{h_{21}}{h_{11}} & \frac{\lvert H \rvert}{h_{11}} \end{pmatrix} = \begin{pmatrix} \frac{D}{B} & -\frac{\lvert T \rvert}{B} \\ -\frac{1}{B} & \frac{A}{B} \end{pmatrix}$
$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}$ $\lvert Z \rvert = z_{11}z_{22} - z_{12}z_{21}$	$\begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} = \begin{pmatrix} \frac{y_{22}}{\lvert Y \rvert} & -\frac{y_{12}}{\lvert Y \rvert} \\ -\frac{y_{21}}{\lvert Y \rvert} & \frac{y_{11}}{\lvert Y \rvert} \end{pmatrix} = \begin{pmatrix} \frac{\lvert H \rvert}{h_{22}} & \frac{h_{12}}{h_{22}} \\ -\frac{h_{21}}{h_{22}} & \frac{1}{h_{22}} \end{pmatrix} = \begin{pmatrix} \frac{A}{C} & -\frac{\lvert T \rvert}{C} \\ \frac{1}{C} & \frac{D}{C} \end{pmatrix}$
$\begin{pmatrix} V_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} I_1 \\ V_2 \end{pmatrix}$ $\lvert H \rvert = h_{11}h_{22} - h_{12}h_{21}$	$\begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} = \begin{pmatrix} \frac{1}{y_{11}} & -\frac{y_{12}}{y_{11}} \\ \frac{y_{21}}{y_{11}} & \frac{\lvert Y \rvert}{y_{11}} \end{pmatrix} = \begin{pmatrix} \frac{\lvert Z \rvert}{z_{22}} & \frac{z_{12}}{z_{22}} \\ -\frac{z_{21}}{z_{22}} & \frac{1}{z_{22}} \end{pmatrix} = \begin{pmatrix} \frac{B}{D} & \frac{\lvert T \rvert}{D} \\ -\frac{1}{D} & \frac{C}{D} \end{pmatrix}$
$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_2 \\ -I_2 \end{pmatrix}$ $\lvert T \rvert = AD - BC$	$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} -\frac{y_{22}}{y_{21}} & -\frac{1}{y_{21}} \\ -\frac{\lvert Y \rvert}{y_{21}} & -\frac{y_{11}}{y_{21}} \end{pmatrix} = \begin{pmatrix} \frac{z_{11}}{z_{21}} & \frac{\lvert Z \rvert}{z_{21}} \\ \frac{1}{z_{21}} & \frac{z_{22}}{z_{21}} \end{pmatrix} = \begin{pmatrix} -\frac{\lvert H \rvert}{h_{21}} & -\frac{h_{11}}{h_{21}} \\ -\frac{h_{22}}{h_{21}} & -\frac{1}{h_{21}} \end{pmatrix}$

This equation describes a procedure that can be used to measure or calculate y_{12}. A short circuit is connected to port 1 to set $V_1 = 0$. A voltage source having voltage V_2 is connected across port 2 and the current, I_1, in the short circuit is calculated. Finally, y_{12} is calculated as the ratio of I_1 to V_2.

Similar procedures can be used to calculate any of the *y*, *z*, *h*, or transmission parameters. Table 3.2 tabulates these procedures.

As an example of how Table 3.2 can be used, consider calculating the *y* parameters of the two port circuit shown in Figure 3.3. Recall that $C_1 = 1$ F, $C_4 = 2$ F, $R_1 = 1/2\ \Omega$, $R_2 = 1/4\ \Omega$, and $R_3 = 1\ \Omega$.) According to Table 3.2, two cases will have to be considered. In the first, a voltage source is connected to port 1 and a short circuit is connected to port 2. In the second, a short circuit is connected across port 1 and a voltage source is connected to port 2. The resulting circuits are shown in Figure 3.5. In Figure 3.5(a)

$$I_1 = \frac{V_1}{R_1 + \frac{1}{C_1 s}} + \frac{V_1}{R_2 + R_3} = \frac{14s + 8}{5s + 10} V_1 = y_{11} V_1 \tag{3.15}$$

and

$$I_2 = -\frac{V_1}{R_2 + R_3} = -\frac{4}{5} V_1 = y_{21} V_1 \tag{3.16}$$

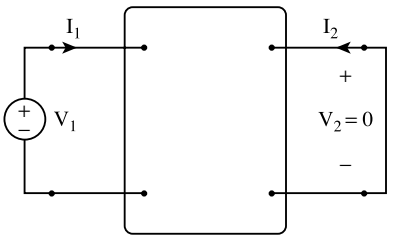
In Figure 3.5(b)

$$I_1 = -\frac{V_2}{R_2 + R_3} = -\frac{4}{5} V_2 = y_{12} V_2 \tag{3.17}$$

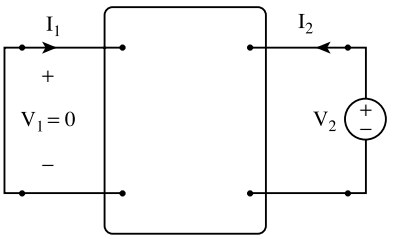
and

$$I_2 = \frac{V_2}{\frac{1}{C_4 s}} + \frac{V_2}{R_2 + R_3} = \frac{10s + 4}{5} V_2 = y_{22} V_2 \tag{3.18}$$

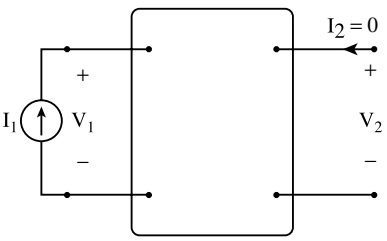
TABLE 3.2 Calculation of Two-Port Parameters



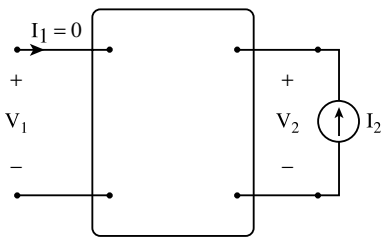
$y_{11} = \dfrac{I_1}{V_1}$ and $y_{21} = \dfrac{I_2}{V_1}$ when $V_2 = 0$



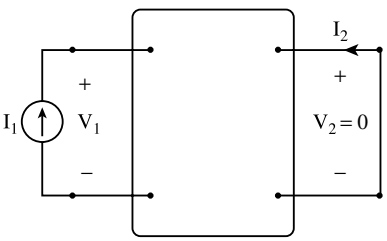
$y_{12} = \dfrac{I_1}{V_2}$ and $y_{22} = \dfrac{I_2}{V_2}$ when $V_1 = 0$



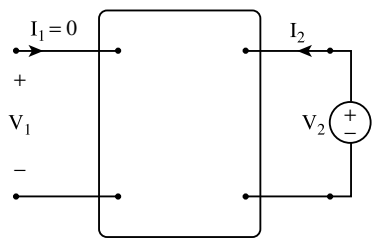
$z_{11} = \dfrac{V_1}{I_1}$ and $z_{21} = \dfrac{V_2}{I_1}$ when $I_2 = 0$



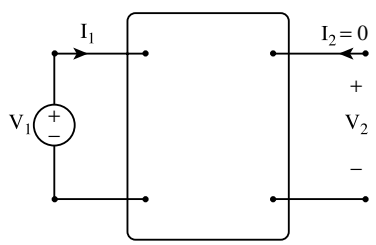
$z_{12} = \dfrac{V_1}{I_2}$ and $y_{22} = \dfrac{V_2}{I_1}$ when $I_1 = 0$



$h_{11} = \dfrac{V_1}{I_1}$ and $h_{21} = \dfrac{I_2}{I_1}$ when $V_2 = 0$

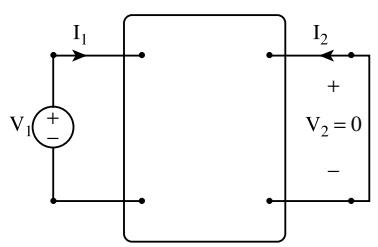


$h_{12} = \dfrac{V_1}{V_2}$ and $h_{22} = \dfrac{I_2}{V_2}$ when $I_1 = 0$



$A = \dfrac{V_1}{V_2}$ when $I_2 = 0$

$C = \dfrac{I_1}{V_2}$ when $I_2 = 0$



$B = \dfrac{V_1}{-I_2}$ when $V_2 = 0$

$D = \dfrac{I_1}{-I_2}$ when $V_2 = 0$

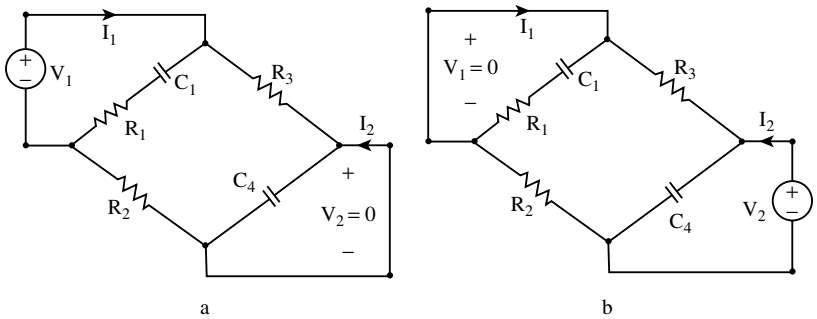


FIGURE 3.5 The test circuits used to calculate the *y* parameters of the example circuit.

Finally, the two-port network is represented by

$$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} \dfrac{14s+8}{5s+10} & -\dfrac{4}{5} \\ -\dfrac{4}{5} & \dfrac{10s+4}{5} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \tag{3.19}$$

To continue the example, row 3 Table 3.1 illustrates how to convert these admittance parameters to hybrid parameters. From Table 3.1,

$$|Y| = \frac{14s+8}{5s+10}\left[\frac{10s+4}{5}\right] - \left(-\frac{4}{5}\right)\left(-\frac{4}{5}\right) = \frac{28s^2+24s}{5(s+2)} \tag{3.20}$$

$$h_{11} = \frac{1}{y_{11}} = \frac{5s+10}{14s+8} \tag{3.21}$$

$$h_{12} = -\frac{y_{12}}{y_{11}} = \frac{2s+4}{7s+4} \tag{3.22}$$

$$h_{21} = \frac{y_{21}}{y_{11}} = -\frac{2s+4}{7s+4} \tag{3.23}$$

$$h_{22} = \frac{|Y|}{y_{11}} = \frac{(14s+12)s}{7s+4} \tag{3.24}$$

To complete the example, notice that Table 3.2 shows how to calculate the hybrid parameters directly from the circuit. According to Table 3.2, h_{22} is calculated by connecting an open circuit across port 1 and a voltage source having voltage V_2 across port 2. The resulting circuit is depicted in Figure 3.6. Now, h_{22} is determined from Figure 3.6(b) by calculating the port current I_2.

$$I_2 = \frac{V_2}{R_1+R_2+R_3+\dfrac{1}{C_1 s}} + \frac{V_2}{\dfrac{1}{C_4 s}} = \frac{14s^2+12s}{7s+4}V_2 = h_{22}V_2 \tag{3.25}$$

Of course, this is the same expression as was calculated earlier from the *y* parameters of the circuit.

Next, consider the problem of analyzing a circuit consisting of a two-port network and some external circuitry. Figure 3.7 depicts such a circuit. The currents in the resistors R_s and R_L are given by

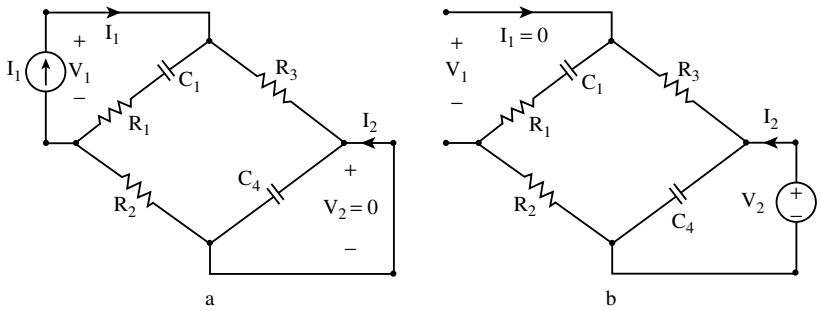


FIGURE 3.6 The test circuits used to calculate the *h* parameters of the example circuit.

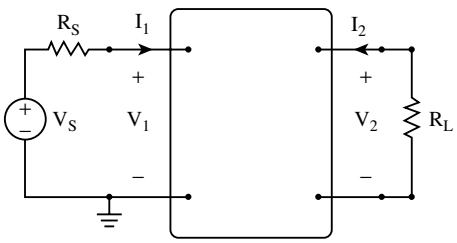


FIGURE 3.7 An application of the two-port network.

$$I_1 = \frac{V_s - V_1}{R_s} \quad \text{and} \quad I_2 = -\frac{V_2}{R_L} \tag{3.26}$$

Suppose the two-port network used in Figure 3.7 is the circuit shown in Figure 3.4 and represented by *y* parameters in Eq. (3.19). Combining the previous expressions for I_1 and I_2 with Eq. (3.19) yields

$$\begin{pmatrix} \frac{V_s}{R_s} \\ 0 \end{pmatrix} = \begin{pmatrix} \frac{14s+8}{5s+10} + \frac{1}{R_s} & -\frac{4}{5} \\ -\frac{4}{5} & \frac{10s+4}{5} + \frac{1}{R_L} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix} \tag{3.27}$$

This equation can then be solved, e.g., using Cramer's Rule, for the transfer function

$$\begin{aligned} \frac{V_2}{V_s}(s) &= \frac{-\left(-\frac{4}{5}\right)\left(\frac{1}{R_s}\right)}{\left(\frac{14s+8}{5s+10} + \frac{1}{R_s}\right)\left(\frac{10s+4}{5} + \frac{1}{R_L}\right) - \left(-\frac{4}{5}\right)^2} \\ &= \frac{4(s+2)R_L}{(28s^2+24s)R_L R_s + (10s+4)(s+2)R_L + (14s+8)R_s + 5(s+2)} \end{aligned} \tag{3.28}$$

Next consider Figure 3.8. This circuit illustrates a caution regarding use of the port convention. The use of ports assumes that the currents in the terminals comprising a port are equal in magnitude and opposite in sign. This assumption is not satisfied in Figure 3.8 so port equations cannot be used to represent the four-terminal network.

Table 3.3 presents three circuit models for two port networks. These three models are based on *y*, *z*, and *h* parameters, respectively. Such models are useful when analyzing circuits that contain subcircuits

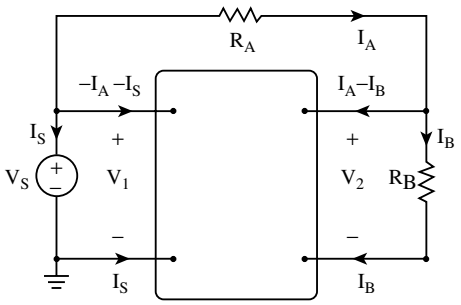


FIGURE 3.8 An incorrect application of the two-port representation.

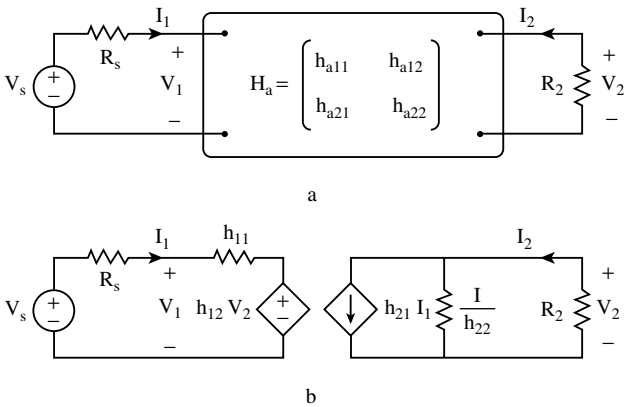


FIGURE 3.9 Application of the circuit model associated with *H* parameters.

that are represented by port parameters. As an example, consider the circuit shown in Figure 3.9(a). This circuit contains a two port network represented by *h* parameters. In Figure 3.9(b) the two-port network has been replaced by the model corresponding to *h* parameters. Analysis of Figure 3.9(b) yields

$$V_s = (R_s + h_{11})I_1 + h_{12}V_2$$
$$V_2 = h_{21}\frac{-R_2}{R_2 h_{22} + 1}I_1 \tag{3.29}$$

After some algebra,

$$\frac{V_2}{V_s} = \frac{h_{21}R_2}{h_{12}h_{21}R_2 - (h_{11} + R_2)(h_{22}R_2 + 1)} \tag{3.30}$$

The circuits in Figure 3.9(a) and (b) are equivalent, so this is the voltage gain of the circuit in Figure 3.9(a).

H parameters are frequently used to describe bipolar transistors. Table 3.4 shows the popular methods for converting this three terminal device into a two-port network. The three configurations shown in Table 3.4 are called "common emitter," "common collector," and "common base" to indicate which terminal is common to both ports. Table 3.4 also presents the notation that is commonly used to name the *h* parameters when they are used to represent a transistor. For example, h_{fe} is the forward gain when the transistor is connected in the common emitter configuration. Comparing with Table 3.3, it can be observed that $h_{fe} = h_{21}$.

TABLE 3.3 Models of Two-Port Networks

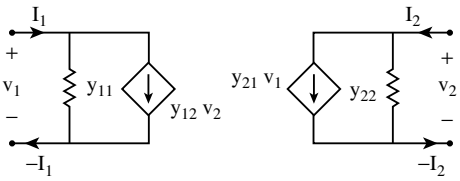
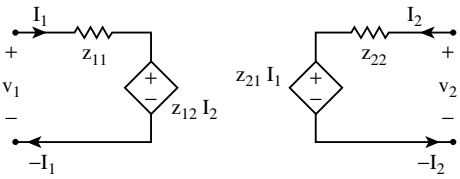
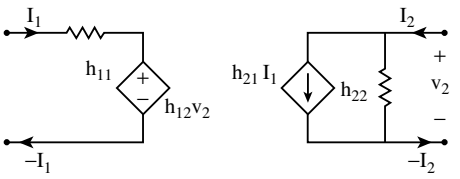
$\begin{pmatrix} I_1 \\ I_2 \end{pmatrix} = \begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} \begin{pmatrix} V_1 \\ V_2 \end{pmatrix}$	
$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} \begin{pmatrix} I_1 \\ I_2 \end{pmatrix}$	
$\begin{pmatrix} V_1 \\ V_2 \end{pmatrix} = \begin{pmatrix} h_{11} & h_{12} \\ h_{21} & h_{22} \end{pmatrix} \begin{pmatrix} I_1 \\ V_2 \end{pmatrix}$	

Figure 3.10 depicts a popular model of a bipolar transistor. Suppose the *h* parameters are calculated for the common emitter configuration. The result is

$$\begin{pmatrix} h_{ie} & h_{re} \\ h_{fe} & h_{oe} \end{pmatrix} = \begin{pmatrix} r_\pi & 0 \\ \beta & r_0 \end{pmatrix} \tag{3.31}$$

This calculation makes a connection between the parameters of the circuit model of the transistor and the *h* parameters used to describe the transistor, e.g., $h_{fe} = \beta$.

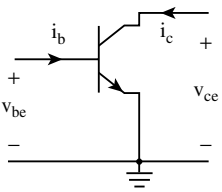
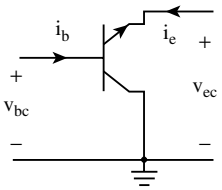
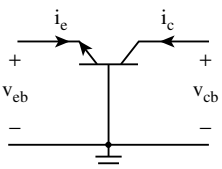
Figures 3.11 to 3.13 illustrate some common interconnections of two-port networks. In Figure 3.11, the two-port network labeled A and B are connected in cascade. As illustrated in Figure 3.11, the transmission matrix of the composite network is given by the product of the transmission matrices of the subnetworks. This simple relationship makes it convenient to use transmission parameters to represent a composite network that consists of the cascade of two subnetworks.

In Figure 3.12, the two-port networks labeled A and B are connected in parallel. As shown in Figure 3.12, the admittance matrix of the composite network is given by the sum of the admittance matrices of the subnetworks. This simple relationship makes it convenient to use admittance parameters to represent a composite network that consists of parallel subnetworks.

In Figure 3.13, the two-port network labeled A and B are connected in series. As illustrated in Figure 3.13, the impedance matrix of the composite network is given by the sum of the impedance matrices of the subnetworks. This simple relationship makes it convenient to use impedance parameters to represent a composite network that consists of series subnetworks.

Figure 3.14 is a circuit that consists of three two-port networks. Two-port parameters representing the entire network can be calculated from the two-port parameters of the subnetworks. Let *c* denote the two-port network consisting of network *b* connected in parallel with network *a*. Represent network *c* with *y* parameters by converting the *h* parameters representing network *a* to *y* parameters and adding these *y* parameters to the *y* parameters representing network *b*.

TABLE 3.4 Using H Parametes to Specify Bipolar Transistors

 common emitter	$\begin{pmatrix} v_{be} \\ i_c \end{pmatrix} = \begin{pmatrix} h_{ie} & h_{re} \\ h_{fe} & h_{oe} \end{pmatrix} \begin{pmatrix} i_b \\ v_{ce} \end{pmatrix}$
 common collector	$\begin{pmatrix} v_{bc} \\ i_e \end{pmatrix} = \begin{pmatrix} h_{ic} & h_{rc} \\ h_{fc} & h_{oc} \end{pmatrix} \begin{pmatrix} i_b \\ v_{ec} \end{pmatrix}$
 common base	$\begin{pmatrix} v_{eb} \\ i_c \end{pmatrix} = \begin{pmatrix} h_{ib} & h_{rc} \\ h_{fb} & h_{oc} \end{pmatrix} \begin{pmatrix} i_b \\ v_{cb} \end{pmatrix}$

i input impedance or admittance
o output impedance or admittance
f forward gain
r reverse gain
e common emitter
b common base
c common collector

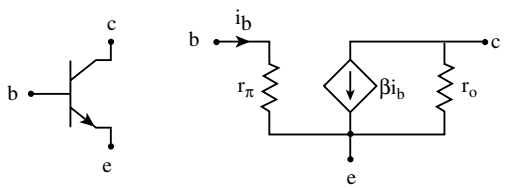
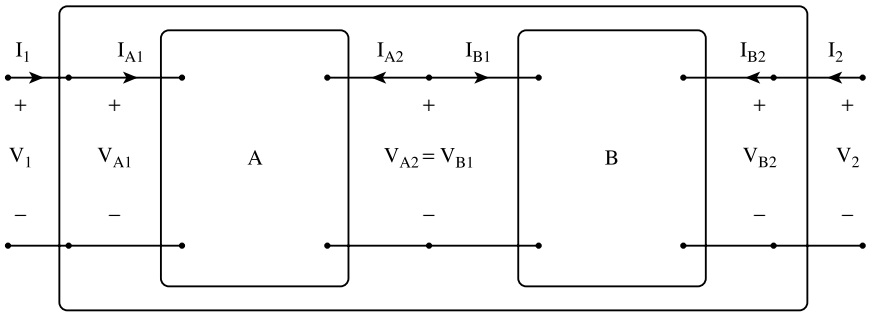


FIGURE 3.10 The hybrid pi model of a transistor.

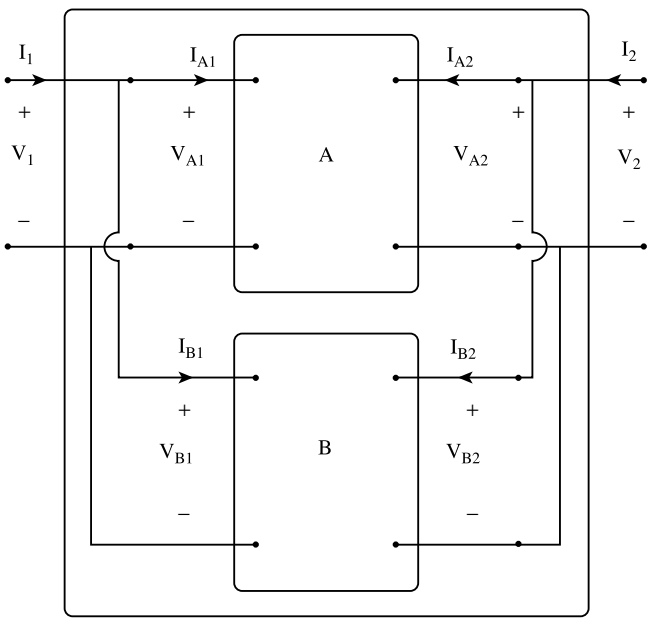
$$Y_c = \begin{pmatrix} \dfrac{1}{h_{a11}} + y_{b11} & -\dfrac{h_{a12}}{h_{a11}} + y_{b12} \\ \dfrac{h_{a21}}{h_{a11}} + y_{b21} & \dfrac{|H_a|}{h_{a11}} + y_{b22} \end{pmatrix} \triangleq \begin{pmatrix} y_{c11} & y_{c12} \\ y_{c21} & y_{c22} \end{pmatrix} \tag{3.32}$$

Next, network *c* is connected in cascade with network *d*. Represent network *d* with transmission parameters by converting the *y* parameters representing network *c* to transmission parameters and multiplying these transmissions parameters by the transmission parameters representing network *d*.



$$\begin{pmatrix} A & B \\ C & D \end{pmatrix} = \begin{pmatrix} A_a & B_a \\ C_a & D_a \end{pmatrix} \begin{pmatrix} A_b & B_b \\ C_b & D_b \end{pmatrix}$$

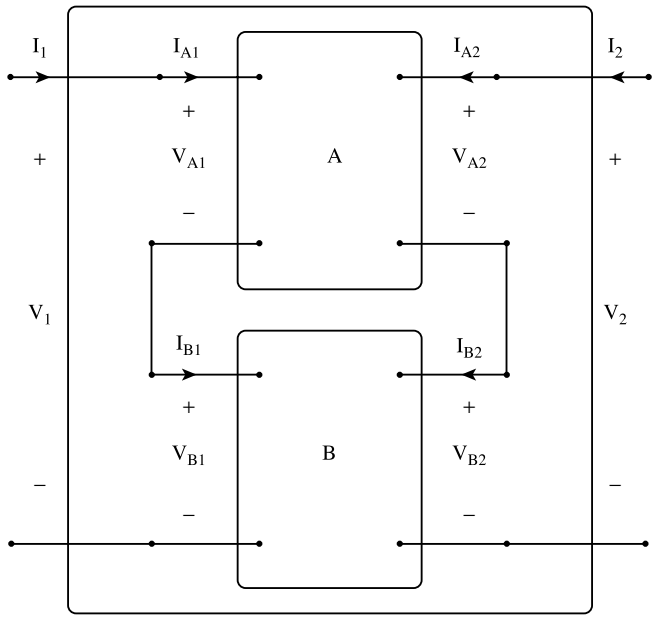
FIGURE 3.11 Cascade connection of two-port networks.



$$\begin{pmatrix} y_{11} & y_{12} \\ y_{21} & y_{22} \end{pmatrix} = \begin{pmatrix} y_{A11} & y_{A12} \\ y_{A21} & y_{A22} \end{pmatrix} + \begin{pmatrix} y_{B11} & y_{B12} \\ y_{B21} & y_{B22} \end{pmatrix}$$

FIGURE 3.12 Parallel connection of two-port networks.

$$T = \begin{pmatrix} -\dfrac{y_{c22}}{y_{c21}} & -\dfrac{1}{y_{c21}} \\ -\dfrac{|Y_c|}{y_{c21}} & -\dfrac{y_{c11}}{y_{c21}} \end{pmatrix} \begin{pmatrix} A_d & B_d \\ C_d & D_d \end{pmatrix} \triangleq \begin{pmatrix} A & B \\ C & D \end{pmatrix} \tag{3.33}$$



$$\begin{pmatrix} z_{11} & z_{12} \\ z_{21} & z_{22} \end{pmatrix} = \begin{pmatrix} z_{A11} & z_{A12} \\ z_{A21} & z_{A22} \end{pmatrix} + \begin{pmatrix} z_{B11} & z_{B12} \\ z_{B21} & z_{B22} \end{pmatrix}$$

FIGURE 3.13 Series connection of two-port networks.

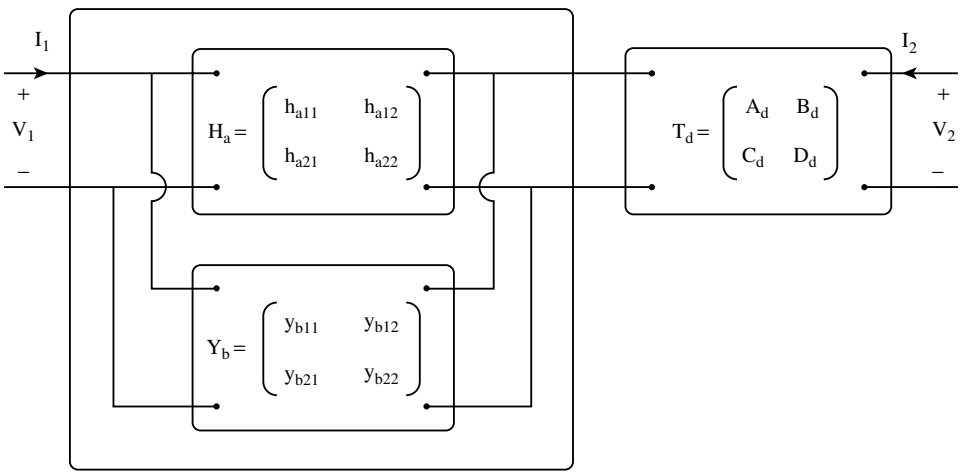


FIGURE 3.14 A circuit consisting of three two-port networks.

Finally, the circuit in Figure 3.14 is represented by

$$\begin{pmatrix} V_1 \\ I_1 \end{pmatrix} = \begin{pmatrix} A & B \\ C & D \end{pmatrix} \begin{pmatrix} V_2 \\ -I_2 \end{pmatrix} \tag{3.34}$$

3.4 Port Representations and Feedback Systems

Port representations of networks can be used to decompose a network into subnetworks. In this section, the decomposition will be done in such a way as to establish a connection between the network and a feedback system represented by the block diagram shown in Figure 3.15. Having established this connection, the theory of feedback systems can be applied to the network. Here, the problem of determining the relative stability of the network will be considered.

The theory of feedback systems [4] can be used to determine relative stability such as the one shown in Figure 3.15. The transfer function of this system is

$$T(\omega) = D(\omega) + \frac{C_1(\omega)A(\omega)C_2(\omega)}{1 + A(\omega)\beta(\omega)} \tag{3.35}$$

The phase and gain margins are parameters of a system that are used to measure relative stability. These parameters can be calculated from $A(\omega)$ and $\beta(\omega)$. To calculate the phase margin, first identify ω_m as the frequency at which

$$\begin{aligned} |A(\omega_m)||\beta(\omega_m)| = 1 \Rightarrow |A(\omega_m)| &= \frac{1}{|\beta(\omega_m)|} \\ \Rightarrow 20\log_{10}|A(\omega_m)| &= -20\log_{10}|\beta(\omega_m)| \end{aligned} \tag{3.36}$$

The phase margin is then given by

$$\begin{aligned} \phi_m &= 180^\circ + \angle\left(A(\omega_m)\beta(\omega_m)\right) \\ &= 180^\circ + \left(\angle A(\omega_m) + \angle\beta(\omega_m)\right) \end{aligned} \tag{3.37}$$

The gain margin is calculated similarly. First, identify ω_p as the frequency at which

$$180^\circ = \angle\left(A(\omega_p)\beta(\omega_p)\right) \tag{3.38}$$

The gain margin is given by

$$GM = \frac{1}{|A(\omega_p)||\beta(\omega_p)|} \tag{3.39}$$

Next, consider an active circuit which can be represented as shown in Figure 3.16. For convenience, it is assumed that the input and output of the circuit are both node voltages. An amplifier has been selected and separated from the rest of the circuit. The rest of the circuit has been denoted as *N*.

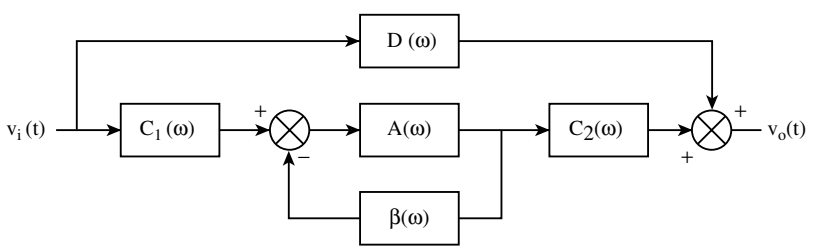


FIGURE 3.15 A feedback system.

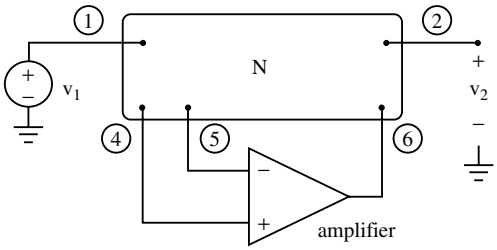


FIGURE 3.16 Identifying the network *N*.

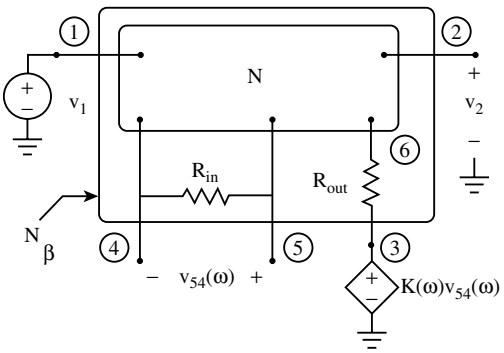


FIGURE 3.17 Identifying the Beta Network, N_β.

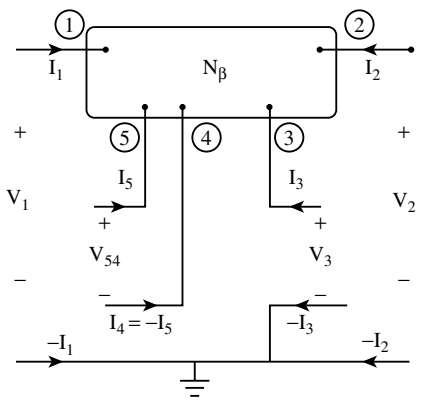


FIGURE 3.18 Identifying the ports of the Beta Network.

Suppose that the amplifier is a voltage controlled voltage source (VCVS), e.g., an inverting or a noninverting amplifier or an op amp. Replacing the VCVS by a simple model yields the circuit shown in Figure 3.17. (The VCVS model accounts for input and output resistance and frequency dependent gain.) Figure 3.17 shows how to identify the network N_β which consists of the network *N* from Figure 3.16 together with the input and output resistances of the VCVS. The network N_β has been called the "Beta Network" [5].

Figure 3.18 illustrates a way of grouping the terminals of the five-terminal network N_β to obtain a four-port network. This four-port network can be represented by the hybrid equation

$$\begin{pmatrix} I_1 \\ V_2 \\ I_3 \\ V_{54} \end{pmatrix} = \begin{pmatrix} H_{11} & H_{12} & H_{13} & H_{14} \\ H_{21} & H_{22} & H_{23} & H_{24} \\ H_{31} & H_{32} & H_{33} & H_{34} \\ H_{41} & H_{42} & H_{43} & H_{44} \end{pmatrix} \begin{pmatrix} V_1 \\ I_2 \\ V_3 \\ I_5 \end{pmatrix} \tag{3.40}$$

Because I_1 and I_3 are not of interest, the first and third rows of this equation can be set aside. Then setting I_2 and I_5 equal to zero yields

$$\begin{pmatrix} V_2 \\ V_{54} \end{pmatrix} = \begin{pmatrix} H_{21} & H_{23} \\ H_{41} & H_{43} \end{pmatrix} \begin{pmatrix} V_1 \\ V_3 \end{pmatrix} \tag{3.41}$$

where, for example

$$H_{43}(\omega) = \frac{V_{54}(\omega)}{V_3(\omega)} \quad \text{when} \quad V_1(\omega) = 0 \tag{3.42}$$

and $H_{21}(\omega)$, $H_{23}(\omega)$ and $H_{41}(\omega)$ are defined similarly.

The amplifier model requires that

$$V_3(\omega) = K(\omega) V_{54}(\omega) \tag{3.43}$$

Combining these equations yields

$$T(\omega) = \frac{V_2(\omega)}{V_1(\omega)} = H_{21}(\omega) + \frac{K(\omega) H_{23}(\omega) H_{41}(\omega)}{1 - K(\omega) H_{43}(\omega)} \tag{3.44}$$

Comparing this equation with the transfer function of the feedback system yields

$$\begin{aligned} A(\omega) &= -K(\omega) \\ \beta(\omega) &= H_{43}(\omega) \\ C_1(\omega) &= H_{23}(\omega) \\ C_2(\omega) &= H_{41}(\omega) \\ D(\omega) &= H_{21}(\omega) \end{aligned} \tag{3.45}$$

These equations establish a correspondence between the feedback system in Figure 3.15 and the active circuit in Figure 3.16. This correspondence can be used to identify $A(s)$ and $\beta(s)$ corresponding to a particular circuit. Once $A(s)$ and $\beta(s)$ are known, the phase or gain margin can be calculated.

Figure 3.19 shows a Tow–Thomas bandpass biquad [3]. When the op amps are ideal devices the transfer function of this circuit is

$$T(s) = \frac{V_{in}(s)}{V_{out}(s)} = \frac{-\frac{1}{KRC} s}{s^2 + \frac{s}{QRC} + \frac{1}{R^2 C^2}} \tag{3.46}$$

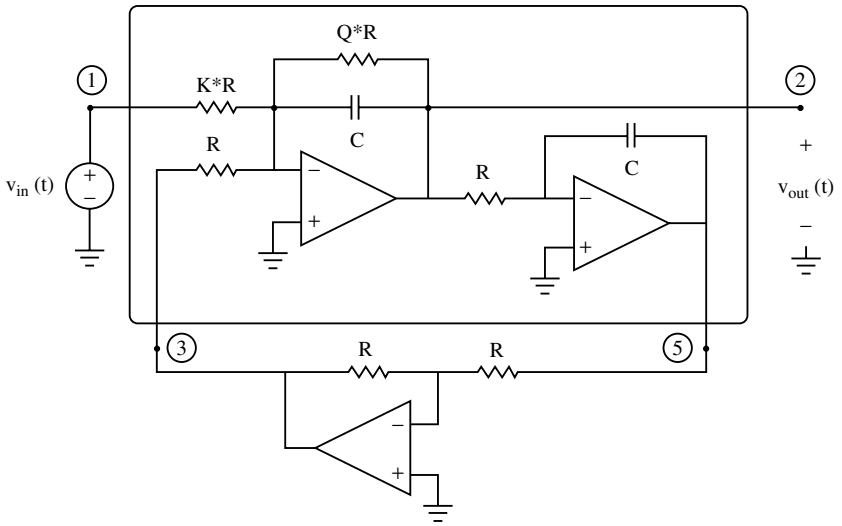


FIGURE 3.19 The Tow–Thomas biquad.

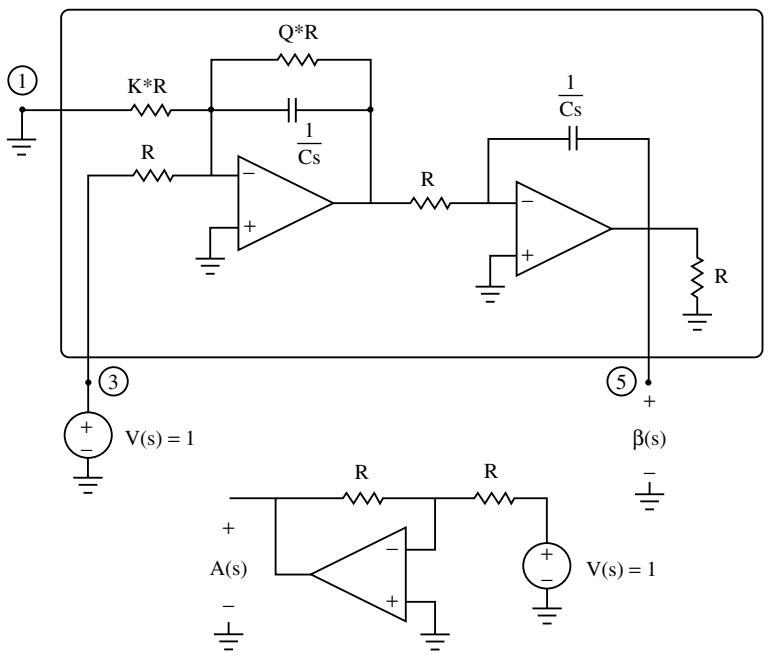


FIGURE 3.20 Calculating $A(s)$ and $\beta(s)$ for the Tow–Thomas biquad.

Figure 3.20 illustrates circuits that can be used to identify $A(s)$ and $\beta(s)$.

$$A(s) = -1$$
$$\beta(s) = \frac{Q}{QR^2C^2s^2 + RCs} \tag{3.47}$$

Next,

$$1 = \left|A(\omega_m)\right|\left|\beta(\omega_m)\right| \Rightarrow (RC\omega_m)^2 = \frac{\frac{1}{Q^2} \pm \sqrt{\frac{1}{Q^4}+4}}{2} \tag{3.48}$$

The phase margin is given by

$$\begin{aligned}\phi_m &= 180^\circ + \angle A(\omega_m) + \angle\beta(\omega_m) \\ &= 180^\circ + 180^\circ - \tan^{-1}\frac{-1}{QRC\omega_m} \\ &= \tan^{-1}\frac{1}{QRC\omega_m}\end{aligned} \tag{3.49}$$

When Q is large,

$$\omega_m \approx \frac{1}{RC} \Rightarrow \phi_m \approx \tan^{-1}\frac{1}{Q} \tag{3.50}$$

When the op amps are not ideal, it is not practical to calculate the phase margin by hand. With computer-aided analysis, accurate amplifier models, such as macromodels, can easily be incorporated into this analysis [5].

3.5 Conclusions

It is frequently useful to decompose a large circuit into subcircuits. These subcircuits are connected together at ports and terminals. Port and terminal parameters describe how the subcircuits interact with other subcircuits but do not describe the inner workings of the subcircuit itself.

This section has presented procedures for determining port and terminal parameters and for analyzing networks consisting of subcircuits which are represented by port or terminal parameters. Port equations were used to establish a connection between electronic circuits and feedback systems.

References

[1] W-K. Chen, *Active Network and Feedback Amplifier Theory*, New York: McGraw-Hill, 1980.
[2] L. P. Huelsman, *Basic Circuit Theory*, Englewood Cliffs, NJ: Prentice Hall, 1991.
[3] L. P. Huelsman and P. E. Allen, *Theory and Design of Active Filters*, New York: McGraw-Hill, 1980.
[4] R. C. Dorf, *Modern Control Systems*, Reading, MA: Addison-Wesley, 1989.
[5] J. A. Svoboda and G. M. Wierzba, "Using PSPICE to determine the relative stability of RC active filters," *Int. J. Electron.*, vol. 74, no. 4, pp. 593–604, 1993.

4

Signal Flow Graphs in Filter Analysis and Synthesis

Pen-Min Lin
Purdue University

4.1 Formulation of Signal Flow Graphs for Linear Networks

Any lumped network obeys three basic laws: Kirchhoff's voltage law (KVL), Kirchhoff's current law (KCL), and the elements' laws (branch characteristics). For filter applications, we write the frequency-domain instead of the time-domain network equations. Three general methods for writing network equations are described in Chapter 2.1. They are the node equations, the loop equations, and the hybrid equations. This section outlines another method, the signal flow graph (SFG) method of characterizing a linear network.

Consider first the construction of signal flow graphs for linear networks without controlled sources. For all practical networks, the independent voltage sources (E) contain no loops, and the independent current sources (J) contain no cutsets. Under these conditions, it is always possible to select a tree T, such that all voltage sources are included in the tree and all current sources are included in the co-tree. The network branches are divided into four sets (each set may be empty) indicated by subscripts as follows:

E: independent voltage sources
J: independent current sources
Z: passive branches in the tree, characterized by impedances
Y: passive branches in the co-tree, characterized by admittances

A step-by-step procedure for constructing an SFG is given below.

Procedure 1 (for linear networks without controlled sources)

Step 1. Apply KVL to express each V_Y (voltage of a passive branch in the co-tree) in terms of V_E and V_Z.
Step 2. Apply KCL to express each I_Z (current of a passive branch in the tree) in terms of I_J and I_Y.
Step 3. For each passive tree branch, consider its voltage as the product of impedance and current, i.e., $V_Z = Z_Z I_Z$.
Step 4. For each passive co-tree branch, consider its current as the product of admittance and voltage, i.e., $I_Y = Y_Y V_Y$.

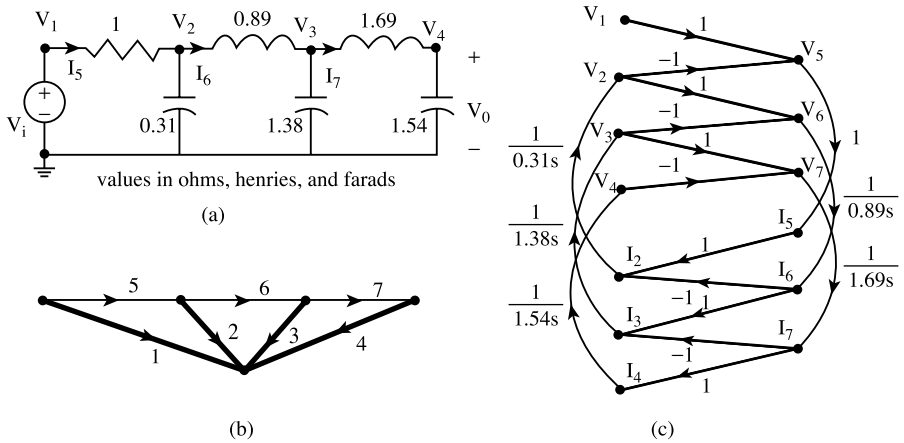


FIGURE 4.1 (a) A low-pass filter network. (b) Directed graph for the network and a chosen tree. (c) SFG based on the chosen tree.

Example 1. Construct a signal flow graph for the low-pass filter network shown in Figure 4.1(a), and use Mason's formula to find the voltage gain function $H(s) = V_o(s)/V_i(s)$.

Solution. The graph associated with the network is shown in Figure 4.1(b) in which the branch numbers and reference directions (passive sign convention) have been assigned. The complexity of the SFG depends on the choice of the tree. In the case of a ladder network, a good tree to use is a star tree which has all tree branches connected to a common node. For the present network, we choose the tree to be $T = \{1, 2, 3, 4\}$, shown in heavy lines in Figure 4.1(b).

Step 1 yields: $V_5 = V_i - V_2, V_6 = V_2 - V_3, V_7 = V_3 - V_4$

Step 2 yields: $I_2 = I_5 - I_6, I_3 = I_6 - I_7, I_4 = I_7$

Step 3 yields: $V_2 = \dfrac{1}{0.31\text{s}} I_2$

$$V_3 = \frac{1}{1.38\text{s}} I_3$$

$$V_4 = \frac{1}{1.54\text{s}} I_4$$

Step 4 yields: $I_5 = V_5$

$$I_6 = \frac{1}{0.89\text{s}} V_6$$

$$I_7 = \frac{1}{1.69\text{s}} V_7$$

The SFG of Figure 4.1(c) displays all the preceding relationships.

Applying Mason's gain formula to the SFG of Figure 4.1(c), we find

$$H(s) = \frac{V_o}{V_i} = \frac{V_4}{V_1} = \frac{1}{s^5 + 3.24s^4 + 5.24s^3 + 5.24s^2 + 3.24s + 1}$$

Next, consider linear networks containing controlled sources. All four types of controlled sources may be present. Our strategy is to utilize procedure 1 described previously with some pre-analysis manipulations. The following is a step-by-step procedure.

Procedure 2 (for linear networks containing controlled sources)

Step 1. *Temporarily* replace each controlled voltage source by an independent voltage source, and each controlled current source by an independent current source, while retaining their original reference directions. The resultant network has no controlled sources.
Step 2. Construct the SFG for the network obtained in step 1 using procedure 1.
Step 3. Express the desired outputs and all controlling variables, if they are not present in the SFG, in terms of the quantities already present in the SFG.
Step 4. Reinstate the constraints of all controlled sources.

Example 2. Construct an SFG for the amplifier circuit depicted in Figure 4.2(a).

Solution. We first replace the controlled voltage source μV_g by an independent voltage source V_x. A tree is chosen to be (V_s, R_a, V_x). The result of step 1 of procedure 2 is depicted in Figure 4.2(b) where dashed lines indicate co-tree branches.

For the links R_b and R_c, we have $I_b = G_b V_b = G_b(V_s - V_a + V_x)$, and $I_c = G_c V_c = -G_c V_x$. For the tree branch R_a we have $V_a = R_a I_a = R_b I_b$. The result of step 2 of procedure 2 is depicted in Figure 4.2(c). Note that the simple relationships $V_b = (V_s - V_a + V_x)$, $V_c = -V_x$ and $I_a = I_b$ have been used to eliminate the variables V_b, V_c and I_a. As a result, these variables do not appear in Figure 4.2(c).

The desired output is $V_o = -V_x$ and the controlling voltage is $V_g = V_s - V_a$. After expressing these relationships in the SFG, step 3 of procedure 2 results in Figure 4.2(d).

Finally, we reinstate the constraint of the controlled source, namely, $V_x = \mu V_g$. The result of step 4 of procedure 2, in Figure 4.2(e), is the desired SFG.

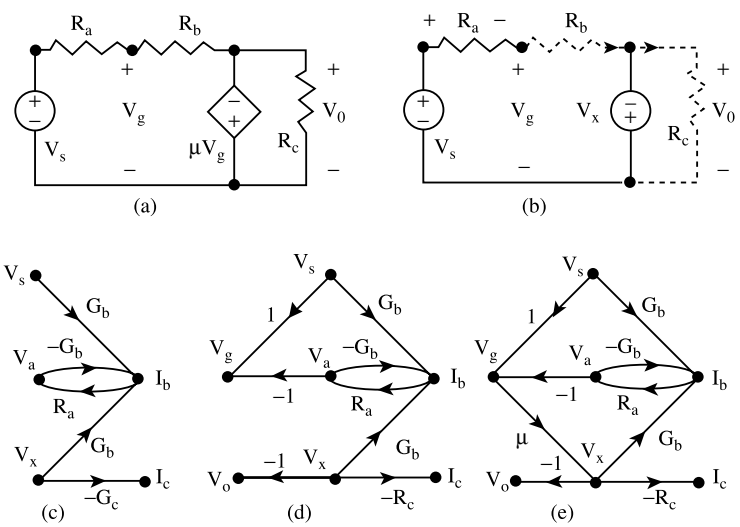


FIGURE 4.2 (a) A linear active network. (b) Result of step 1, procedure 2. (c) Result of step 2, procedure 2. (d) Result of step 3, procedure 2. (e) The desired SFG.

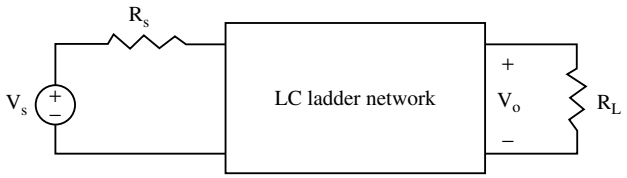


FIGURE 4.3 A doubly terminated passive filter.

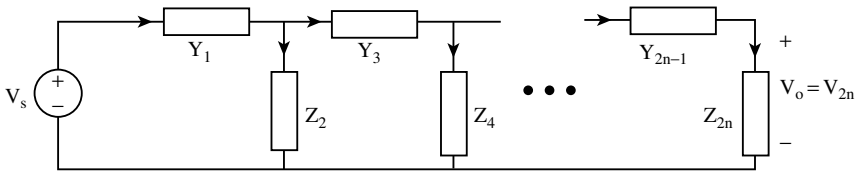


FIGURE 4.4 A general ladder network.

4.2 Synthesis of Active Filters Based on Signal Flow Graph Associated with a Passive Filter Circuit

The preceding section demonstrates that the equations governing a linear network can be described by an SFG in which the branch weights (or transmittances) are either real constants or simple expressions of the form K_s or K/s. All the cause–effect relationships displayed in such an SFG can, in turn, be implemented with resistors, capacitors, and ideal operational amplifiers. The inductors are *not* needed in the implementation. Whatever frequency response prevailing in the original linear circuit appears exactly in the RC-op-amp circuit.

In active filter synthesis, the method described in Section 4.1 is applied to a passive filter in the form of an LC (inductor-capacitor) ladder network terminated in a resistance at both ends as illustrated in Figure 4.3. The reason is that this type of filter, with $R_s = R_L$, has been proved to have the best sensitivity property [1, p. 196]. By this, we mean that the frequency response is least sensitive with respect to the changes in element values, when compared to other types of filter circuits. Because magnitude scaling (i.e., multiplying all impedances in the network by a factor K_m) does not affect the voltage gain function, we always normalize the prototype passive filter network so that the source resistance becomes 1 Ω. The advantage of this normalization will become evident in several examples given in this section.

The SFG illustrated in Figure 4.1(c) has many branches crossing each other. For a ladder network, with a proper choice of the tree and a rearrangement of the SFG nodes, all crossings can be eliminated. To achieve this, we first label a general ladder network as shown in Figure 4.4.

The following conventions are used in the labels of Figure 4.4:

1. All series branches are numbered odd and characterized by their admittances.
2. All shunt branches are numbered even and characterized by their impedances.
3. A single arrow is used to indicate the reference directions of both the voltage and the current of each network branch. Passive sign convention is used.
4. If the LC ladder in Figure 4.4 has a series element at the source end, then Y_1 represents that element in series with R_s.
5. If the LC ladder in Figure 4.4 has a shunt element at the load end, then Z_{2n} represents that element in parallel with R_L.

For constructing the SFG, choose a tree to consist of the voltage source and all shunt branches. The SFG for the circuit may be constructed using procedure 1 of Section 4.1. First, list the equations obtained in each step.

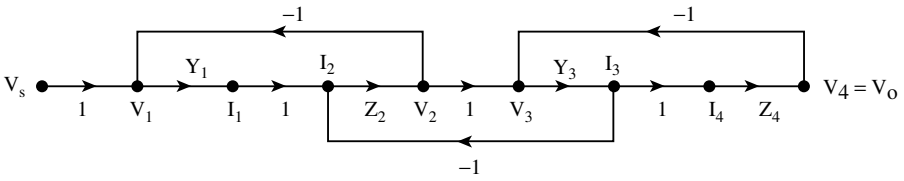


FIGURE 4.5 SFG for a 4-element ladder network.

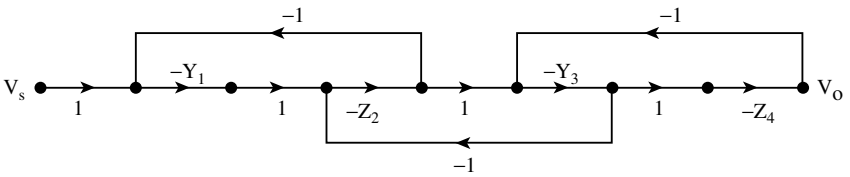


FIGURE 4.6 Inverting integrators are used in this modified SFG.

$$\text{Step 1.} \quad V_1 = V_s - V_2, \;\; V_3 = V_2 - V_4, \ldots, V_{2n-1} = V_{2n-2} - V_{2n}$$

$$\text{Step 2.} \quad I_2 = I_1 - I_3, \;\; I_4 = I_3 - I_5, \ldots, I_{2n} = I_{2n-1}$$

$$\text{Step 3.} \quad V_2 = Z_2 I_2, \;\; V_4 = Z_4 I_4, \ldots, V_{2n} = Z_{2n} I_{2n}$$

$$\text{Step 4.} \quad I_1 = Y_1 V_1, \;\; I_3 = Y_3 V_3, \ldots, I_{2n-1} = Y_{2n-1} V_{2n-1}$$

These relationships are represented by the SFG in Figure 4.5 for the case of a four-element ladder network. Note that the SFG graph nodes have been arranged in such a way that there are no branch crossings. The pattern displayed in this SFG suggests the children's game of *leapfrog*. Consequently, an active filter synthesis based on the SFG of Figure 4.5 is called a leapfrog realization. The transmittance of each SFG branch indicates the type of mathematical operation performed. For example, $1/s$ means integration and is implemented by an op amp integrator. Likewise, $1/(s + a)$ is implemented by a lossy op amp integrator. It is well known that inverting integrators and inverting summers can be designed with singled-ended op amps (i.e., the noninverting input terminal of each op amp is grounded), [2–5]. Noninverting integrators and noninverting summers can also be designed, but require differential-input op amps and more complex circuitry. Therefore, there is an advantage in using the inverting types. To this end, we multiply all Z's and Y's in Figure 4.5 by –1, with the result shown in Figure 4.6. Note that in Figure 4.6 we have removed the labels of internal SFG nodes because they are of no consequence in determining the transfer function. The transfer function V_o/V_s is the same for both Figure 4.5 and Figure 4.6. This is quite obvious from Mason's gain formula, as all path weights and loop weights are not affected by the modification. A branch transmittance of –1 indicates an inverting amplifier. In the interest of reducing the total number of op amps used, we want to reduce the number of branches in the SFG that have weight –1. This can be achieved by inserting branches weighted –1 in some strategic places. Each insertion will lead to the change of the signs of one or two feedback branches. The rules are (a) inserting a branch weighted –1 in a forward path segment shared by two feedback loops changes the signs of the two feedback branch weights; (b) inserting a branch weighted –1 in a forward path segment belonging to one feedback loop only changes the signs of that feedback branch weight. Figure 4.6 is modified this way and the result is shown in Figure 4.7. The inserted branches are shown in heavy lines.

Comparing Figure 4.6 with Figure 4.7, we see that there is no change in path weights and loop weights. Therefore, Mason's gain formula assures that both SFG have the same transfer function. For a six-element ladder network, three branches weighted –1 must be inserted. This leads to a sign change of the single forward path weight in the SFG, and the output node variable now becomes $-V_o$. For filter applications

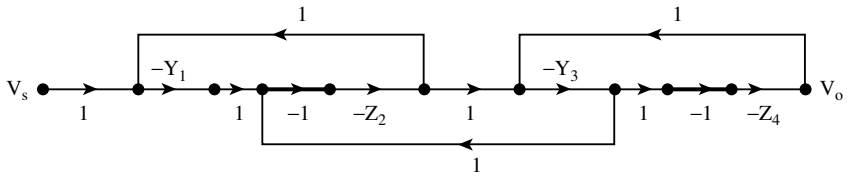
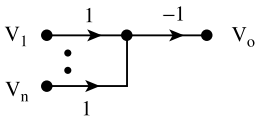
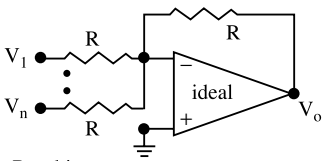
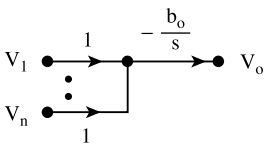
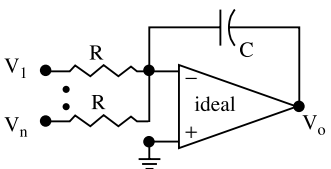
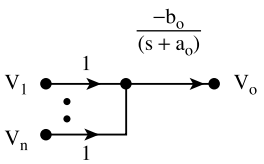
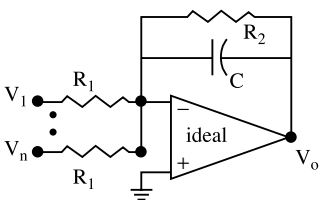


FIGURE 4.7 Modification to reduce the number of inverting amplifiers.

TABLE 4.1 Component Op Amp Circuits for Synthesizing Active Low-Pass Filters by the Leapfrog Technique

(1)	 $V_o = -(V_1 + \ldots + V_n)$	Inverting summer  R: arbitrary
(2)	 $V_o = -\frac{b_o}{s}(V_1 + \ldots + V_n)$	Inverting summing integrator  C: arbitrary, $R = \frac{1}{b_o C}$
(3)	 $V_o = \frac{-b_o}{(s + a_o)} \quad (V_1 + \ldots + V_n)$	Inverting summing lossy integrator  C: arbitrary, $R_1 = \frac{1}{b_o C}$, $R_2 = \frac{1}{a_o C}$,

Note: Each component RC-op-amp circuit in the right column may be magnitude-scaled by an arbitrary factor.

this change of sign in the transfer function is acceptable as we are concerned mainly with the magnitude response.

An implementation of the SFG of Figure 4.7 may be accomplished easily by referring to Table 4.1 and picking the component op amp circuits for realizing the SFG transmittances -1, $-Y_1$, $-Z_2$, etc. Figure 4.7 dictates how these component op amp circuits are interconnected to produce the desired voltage gain function. An example will illustrate the procedure.

Example 3. Figure 4.8 shows a normalized Butterworth fourth-order, low-pass filter, where the 1-Ω source resistance has been included in Y_1, and the 1-Ω load resistance included in Z_4.

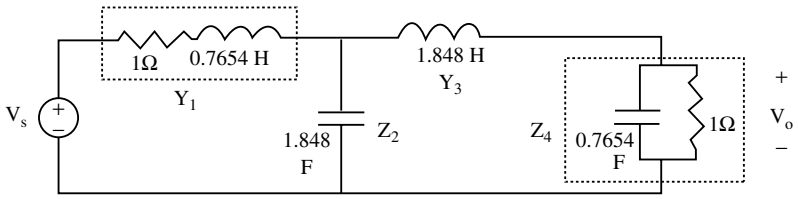


FIGURE 4.8 A fourth-order, Butterworth low-pass filter.

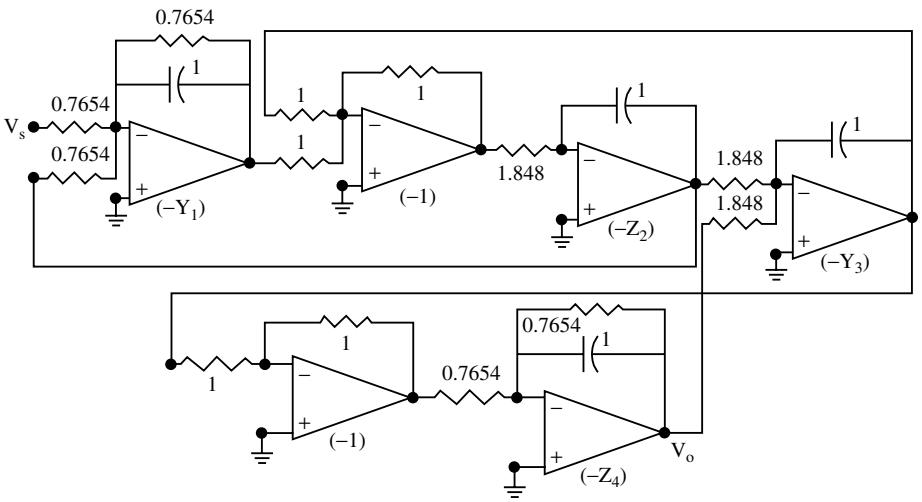


FIGURE 4.9 Leapfrog realization of passive filter of Figure 4.8. For the normalized case of $\omega_{3dB} = 1$ r/sec, values are in Ω and F. For a practical case of $\omega_{3dB} = 10^6$ rad/s, values are in kΩ and nF.

The leapfrog-type SFG for this circuit, after suitable modifications, is shown in Figure 4.7, where

$$-Y_1 = -\frac{1}{0.7654s + 1}$$

$$-Z_2 = -\frac{1}{1.848s}$$

$$-Y_3 = -\frac{1}{1.848s}$$

$$-Z_4 = -\frac{1}{0.7654s + 1}$$

The SFG branch transmittance $-Z_2$ and $-Y_3$ are realized using item (2) of Table 4.1, while $-Y_1$ and $-Z_4$ use item (3). The two SFG branches with weight –1 in Figure 4.7 require item (1). The SFG branches with weight 1 merely indicate how to feed the inputs to each component network. No additional op amps are needed for such SFG branches. Thus, a total of six op amps are required. The interconnection of these component circuits is described by Figure 4.7. The complete circuit is shown in Figure 4.9. One-farad capacitances have been used in the circuit. Recall that the original passive low-pass filter has a 3-dB frequency of 1 rad/s. By suitable magnitude scaling and frequency scaling, all element values in the active filter of Figure 4.9 can be made practical. For example, if the 3-dB frequency is changed to 10^6 rad/s, then the capacitances in Figure 4.9 are divided by 10^6. We may arbitrarily magnitude scale the resultant circuit by a factor of 10^3. Then, all resistances are multiplied by 10^3 and all capacitances are

further divided by 10^3. The final circuit is still the one shown in Figure 4.9, but with resistances in kΩ and capacitance in nF. The parenthetical quantity beside each op amp indicates the type of transfer function it produces.

If a doubly terminated passive filter has a shunt reactance at the source end and a series reactance the load end, then its dual network has a series reactance at the source end and a shunt reactance at the load end. The voltage gain functions of the original network and its dual differ at most by a multiplying constant. We can apply the method to the dual network.

For doubly terminated Butterworth and Chebyshev low-pass filters of odd orders, the passive filter either has series reactances or shunt reactances at both ends. The next example shows the additional SFG manipulations needed to construct the RC-op-amp circuit.

Example 4. Obtain a leapfrog realization of the third-order Butterworth low-pass filter shown in Figure 4.10(a).

Solution. The network is again a four-element ladder network with a modified SFG in terms of the series admittances and shunt impedances as depicted in Figure 4.6. Note that the 1-Ω source resistance alone constitutes the element Y_1. Inserting a branch weighted –1 in front of $-Y_3$ changes the weights of two feedback branches from –1 to 1, and the output from V_o to $-V_o$. Figure 4.10(b) gives the result. Next, apply the node absorption rule to remove nodes V_A and V_B in Figure 4.10b. The result is Figure 4.10(c).

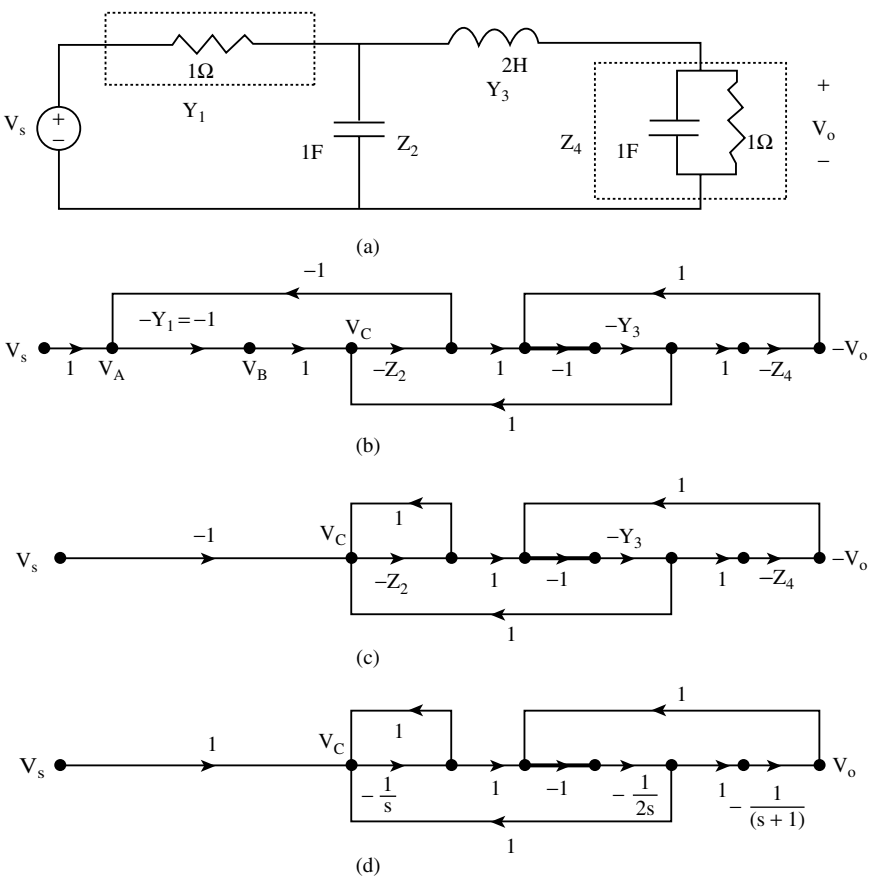


FIGURE 4.10 Leapfrog realization of a third-order, Butterworth low-pass filter. (a) The passive prototype. (b) Leapfrog SFG simulation. (c) Absorption of SFG nodes. (d) Final SFG for active filter realization.

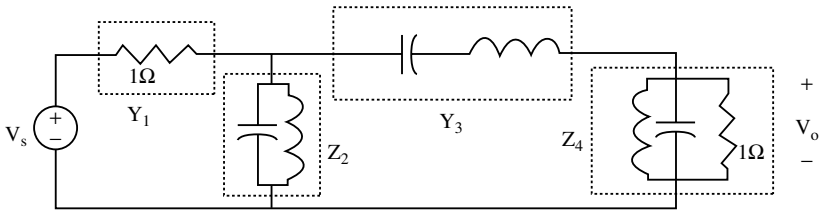


FIGURE 4.11 A bandpass passive filter derived from the circuit of Figure 4.10(a).

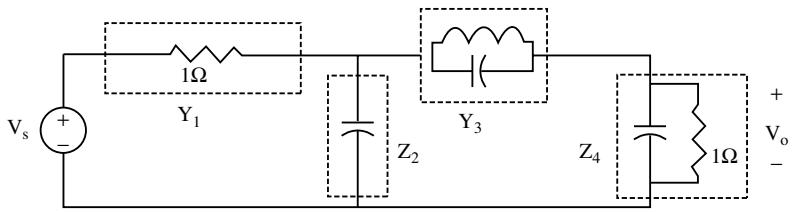


FIGURE 4.12 Network configuration of a doubly terminated filter having a third-order, elliptic or inverse Chebyshev low-pass response.

Finally, we recognize that the left-most branch weight –1 is not contained in any loop weights, and appears in the single forward path weight. Therefore, if this branch weight is changed from –1 to 1, the output will be changed from $-V_o$ to V_o. When this change is made, and all specific branch weights are used, the final SFG is given in Figure 4.10(d). The circuit implementation is now a simple matter of picking component networks from Table 4.1 and connecting them as in Figure 4.10(d). A total of four op amps are required, one each for the branch transmittance $-1/s$, $-1/(2s)$, $-1/(s+1)$, and -1.

Passive bandpass filters may be derived from low-pass filters using the frequency transformation technique described in Chapter 72. The configuration of a bandpass filter derived from the third-order Butterworth filter of Figure 4.10(a) is given in Figure 4.11.

The impedance and admittance functions Z_2, Y_3, and Z_4 are of the form

$$\frac{s}{a_2 s^2 + a_1 s + a_0}$$

The SFG thus contains quadratic branch transmittances. Several single op amp realizations of the quadratic transmittances are discussed in Chapter 82, while some multiple op amp realizations are presented in the next subsection. The interconnection of the component networks, however, is completely specified by an SFG similar to Figure 4.7 or Figure 4.10(d). Complete design examples of this type of bandpass active filter may be found in many books [2–5].

The previous example shows the application of the leapfrog technique to low-pass and bandpass filters of the Butterworth or Chebyshev types. The technique, when applied to a low-pass filter having an elliptic response or an inverse Chebyshev response will require the use of some differentiators. The configuration of a third order low-pass elliptic filter or inverse Chebyshev filter is depicted in Figure 4.12. Notice that Y_3 has the expression

$$Y_3 = a_2 s + \frac{a_0}{s} = \frac{a_2 s^2 + a_0}{s}$$

The term $a_2 s$ in the voltage gain function of the component network clearly indicates the need of a differentiator. An example of such a design may be found in [1, pp. 382–385].

As a final point in the leapfrog technique, consider the problem of impedance normalization. In all the previous examples, the passive prototype filter has equal terminations and has been magnitude-scaled so that $R_s = R_L = 1$. Situations occur where the passive filter has unequal terminations. For example, the passive filter may have $R_s = 100\ \Omega$ and $R_L = 400\ \Omega$ in a four-element ladder network in Figure 4.8. Three possibilities will be considered.

(1) No impedance normalization is done on the passive filter. Then,

$$-Y_1 = -\frac{1}{L_1 s + 100}$$

$$-Z_4 = -\frac{1}{C_4 s + \dfrac{1}{400}}$$

From Table 4.1, the lossy integrator realizing $-Y_1$ has a resistance ratio of 100, and the resistance ratio for the $-Z_4$ circuit is 400. Such a large ratio is undesirable.

(2) An impedance normalization is done with $R_o = 100\ \Omega$ so that R_s becomes 1 and R_L becomes 4. Then

$$-Y_1 = -\frac{1}{L_1 s + 1}$$

$$-Z_4 = -\frac{1}{C_4 s + \dfrac{1}{4}}$$

The resistance ratio in the lossy integrator now becomes 1 for the $-Y_1$ circuit, and 4 for the $-Z_4$ circuit — an obvious improvement over the non-normalized case.

(3) An impedance normalization is done with $R_o = \sqrt{R_s R_L} = 200$. Then $R_s = 0.5$, $R_L = 2$, and

$$-Y_1 = -\frac{1}{L_1 s + 0.5}$$

$$-Z_4 = -\frac{1}{C_4 s + 0.5}$$

The resistance ratio in the lossy integrator is now 2 for both the $-Y_1$ circuit and the $-Z_4$ circuit — a further improvement over case (2) using $R_o = R_s$.

The conclusion is that, in the interest of reducing the spread of resistance values, the best choice of R_o for normalizing the passive filter is $R_o = \sqrt{R_s R_L}$. For the case of equal terminations, this choice leads to $R_s = R_L = 1$.

Instead of starting with a normalized passive filter, one can also construct a leapfrog-type SFG based on the unnormalized passive filter. For a four-element ladder network, the result is given in Figure 4.7. We now perform the following SFG manipulation, which has the same effect as the impedance normalization of the passive filter: Select a normalization resistance, R_o and divide all Z's in the SFG by R_o, and multiply all Y's by R_o. The resultant SFG is given in Figure 4.13.

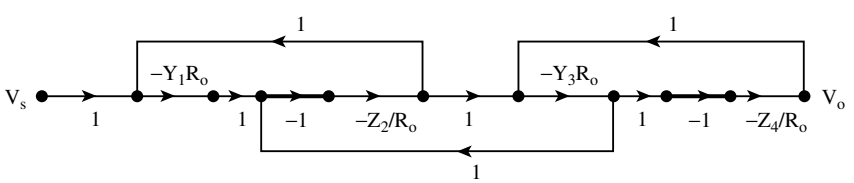


FIGURE 4.13 Result of normalization of the SFG of Figure 4.7.

It is easy to see that the SFG in both Figures 4.7 and 4.13 have the same loop weights and single forward path weight. Therefore, the voltage gain function remains unchanged with the normalization process. One advantage of using the normalized SFG is that the branch transmittances $Y_k R_o$ and Z_k/R_o are dimensionless, and truly represent voltage gain function of component op amp circuits [2, p. 288].

4.3 Synthesis of Active Filters Based on Signal Flow Graph Associated with a Filter Transfer Function

The preceding section describes one application of the SFG in the synthesis of active filters. The starting point is a passive filter in the form of a doubly terminated LC ladder network. In this section, we describe another way of using the SFG technique to synthesize an active filter. The starting point in this case is a filter transfer function instead of a passive network.

Let the transfer voltage ratio function of a filter be

$$\frac{V_o}{V_i} = H(s) = \frac{b_m s^m + b_{m-1} s^{m-1} + \ldots + b_1 s + b_o}{s^n + a_{n-1} s^{n-1} + \ldots + a_1 s + a_0}, \quad m \le n \tag{4.1}$$

By properly selecting the coefficient a's and b's, all types of filter characteristics can be obtained: low-pass, high-pass, bandpass, band elimination, and all-pass. We assume that these coefficients have been determined. Our problem is how to realize the transfer function using SFG theory and RC-op-amp circuits.

For the present application, we impose two constraints on the signal flow graph:

1. No second-order or higher-order loops are present. In other words, all loops in the SFG touch each other.
2. Every forward path from the source node to the output node touches all loops.

For such a special kind of SFG, Mason's gain formula reduces to

$$\frac{V_o}{V_i} = H(s) = \frac{\sum P_k}{1 - (L_1 + L_2 + \ldots + L_n)} \tag{4.2}$$

where L_n is the nth loop weight, P_k is the kth forward path weight, and summations are over all forward paths and all loops. Our strategy is to manipulate Eq. (4.1) into the form of Eq. (4.2), and then construct an SFG to have the desired loops and paths, meeting constraints (1) and (2). Integrators are preferred over differentiators in actual circuit implementation, therefore, we want $1/s$ instead of s to appear as the SFG branch transmittances. This suggests the division of both the numerator and denominator of Eq. (4.1) by s^n, the highest degree term in the denominator.

The result is

$$\begin{aligned} \frac{V_o}{V_i} &= H(s) \\ &= \frac{b_m \left(\frac{1}{s}\right)^{n-m} + b_{m-1} s^{n-m+1} + \ldots + b_1 \left(\frac{1}{s}\right)^{n-1} + b_o \left(\frac{1}{s}\right)^n}{1 + a_{n-1}\left(\frac{1}{s}\right) + \ldots + a_1 \left(\frac{1}{s}\right)^{n-1} + a_0 \left(\frac{1}{s}\right)^n}, \quad m \le n \end{aligned} \tag{4.3}$$

Comparing Eq. (4.3) with Eq. (4.2), we can identify the loop weights

$$
\begin{aligned}
L_1 &= -a_{n-1}(1/s) \\
L_2 &= -a_{n-2}(1/s)^2 \\
&\cdots \\
L_n &= -a_o(1/s)^n
\end{aligned}
\tag{4.4}
$$

and the forward path weights

$$
b_m\left(\frac{1}{s}\right)^{n-m}, \quad b_{m-1}\left(\frac{1}{s}\right)^{n-m+1}, \quad \ldots b_1\left(\frac{1}{s}\right)^{n-1}, \quad b_o\left(\frac{1}{s}\right)^n \tag{4.5}
$$

Many SFGs may be constructed to have such loop and path weights, and the touching properties stated previously in (1) and (2). Two simple ones are given in Figure 4.14(a) and (b) for the case $n = m = 3$. The extension to higher-order transfer functions is obvious. In control theory, the system represented by Figure 4.14(a) is said to be of the controllable canonical form, and Figure 4.14(b) the observable canonical form. In a filter application, we need not be concerned about the controllability and observability of the system. The terms are used here merely for the purpose of circuit identification. Our major concern here is how to implement the SFG by an RC-op-amp circuit.

An SFG branch having transmittance $1/s$ indicates an integrator. If the terminating node of the $1/s$ branch has no other incoming branches [as in Figure 4.14(a)], then that node variable represents the output of the integrator. On the other hand, if $1/s$ is the transmittance of only one of several incoming branches incident at the node V_k [as in Figure 4.14(b)], then V_k is *not* the output of an integrator. In order to identify the integrator outputs clearly for the purpose of circuit interconnection, we insert some dummy branches with weight 1 in series with the branches weighted $1/s$. When this is done to Figure 4.14(b), the result is Figure 4.15 with the inserted dummy branches shown in heavy lines. An SFG branch with weight $-1/s$ represents an inverting integrator. As pointed out in Section 4.2, the circuitry of an inverting integrator is simpler than that of a noninverting integrator. To have an implementation utilizing inverting integrators, we replace all SFG branch weights $1/s$ in Figure 4.14 by $-1/s$. In order to

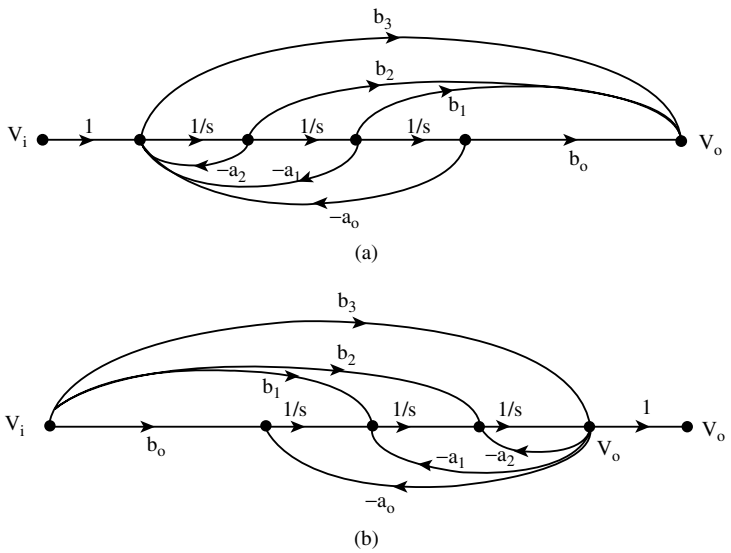


FIGURE 4.14 Two simple SFGs having a gain function given by Eq. (4.3). (a) Controllable canonical form. (b) Observable canonical form.

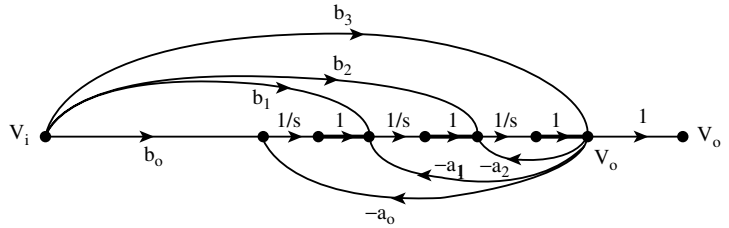


FIGURE 4.15 Insertion of dummy branches to identify integrator outputs.

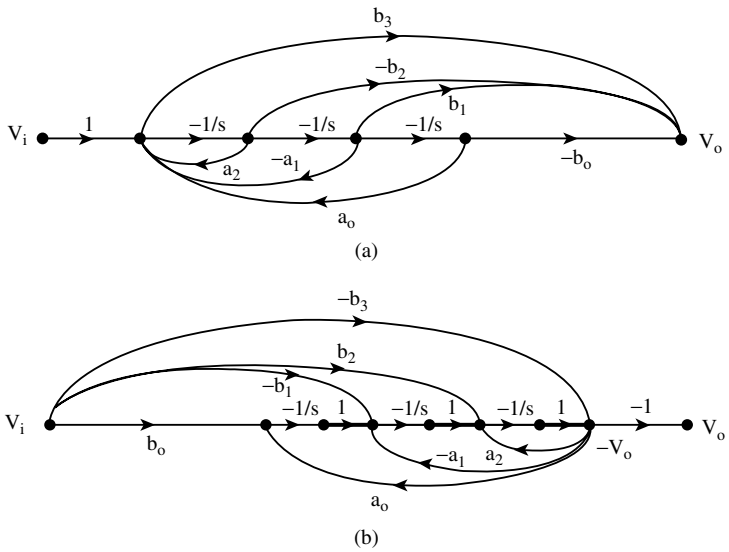


FIGURE 4.16 Simulation of $H(s)$ by an SFG containing inverting integrators.

maintain the original path and loop weights, the signs of some feedback branches and forward path branches must be changed accordingly. When this is done, Figure 4.14(a) and Figure 4.15 become those shown in Figure 4.16(a) and (b), respectively. Our next goal is to implement these SFGs by RC-op-amp circuits. Because SFGs of the kind described in this section are widely used in the study of linear systems by the state variable approach, the active filters based such SFGs are called *state variable filters* [6].

Example 5. Synthesize a state variable active filter to have a third order Butterworth low-pass response having 3 dB frequency $\omega_o = 10^6$ rad/s. All op amps used are single-ended.

Solution. As usual in filter synthesis, we first construct the filter for the normalized case, i.e., $\omega_o = 1$ rad/s, and then perform frequency scaling to obtain the required filter. The normalized voltage gain function of the filter is

$$H(s) = \frac{V_o}{V_i} = \frac{1}{s^3 + 2s^2 + 2s + 1} \tag{4.6}$$

and the two SFGs in Figure 4.16 become those depicted in Figure 4.17.

Because we are concerned with the magnitude response only, $-V_o$ instead of V_o can be accepted as the desired output. Therefore, in Figure 4.17, the rightmost SFG branch with gain (–1) need not be implemented. The implementation of the SFG as RC-op-amp circuits is now just a matter of looking up Table 4.2, selecting proper component networks and connecting them as specified by Figure 4.17. The

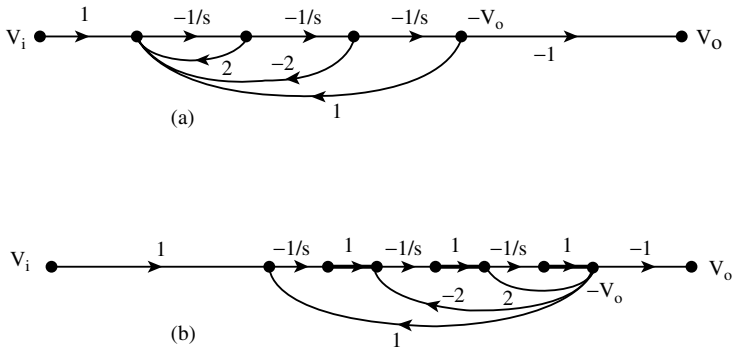
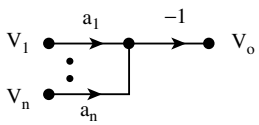
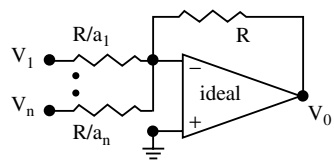
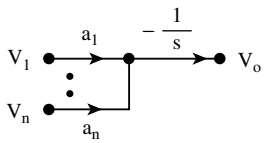
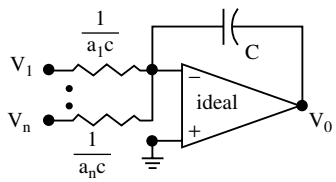
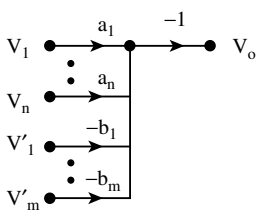
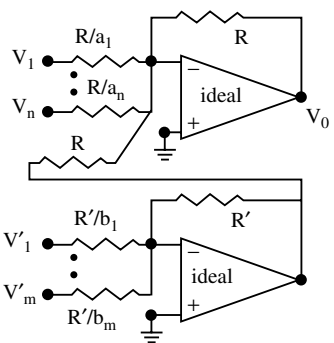


FIGURE 4.17 Two SFG representations of Eq. (4.6).

TABLE 4.2 Single-Ended Op Amp Circuits for Implementing State Variable Active Filters

Signal flow graph	RC-op-amp circuit
(1)  $V_o = -(a_1V_1 + \ldots + a_nV_n)$	Inverting scaled summer  R: arbitrary
(2)  $V_o = -\frac{1}{s}(a_1V_1 + \ldots + a_nV_n)$	Inverting scaled summing integrator  C: arbitrary
(3)  $V_o = -(a_1V_1 + \ldots + a_nV_n) + (b_1V'_1 + \ldots + b_mV_n)$	Bi-polarity summer  R and R′: arbitrary

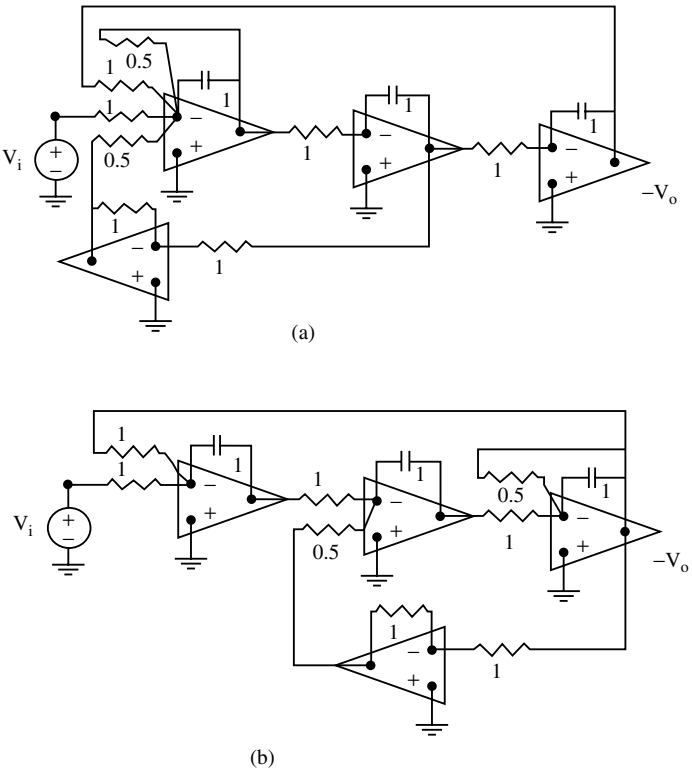


FIGURE 4.18 Two op amp circuit realizations of $H(s)$ given by Eq. (4.6).

results are given in Figure 4.18(a) and (b). These circuits, with element values in ohms and farads, realize the normalized transfer function having $\omega_c = 1$ rad/s. To meet the original specification of $\omega_c = 10^6$ rad/s, we frequency-scale the circuits by a factor 10^6 (i.e., divide all capacitances by 10^6). To have practical resistance values, we further magnitude-scale the circuits by a factor of, say, 1000. The resistances are multiplied by 1000, and the capacitances are further divided by 1000. The final circuits are still given by Figure 4.18, but now with element values in kΩ and nF.

In example 5, both realizations require 4 op amps. In general, for an nth order transfer function given by Eq. (4.1) with all coefficients nonzero, a synthesis based on Figure 4.16(a) (controllable canonical form) requires $n + 3$ single-ended op amps. The breakdown is as follows [refer to Figure 4.16(a)]:

n inverting scaled integrators (item 2, Table 4.2) for the n SFG branches with weight $-1/s$
2 op amps for the bipolarity summer (item 3, Table 4.2) to obtain V_o
1 inverting scaled summer (item 1, Table 4.2) to invert and add up signals from branches with weights $-a_1$, $-a_3$, etc., before applying to the left-most integrator

On the other hand, a synthesis based on Figure 4.16(b) (observable canonical form) requires only $n + 2$ single-ended op amps. To see this, we redraw Figure 4.16(b) as Figure 4.19 by inserting branches with weight –1, and making all literal coefficients positive. The breakdown is as follows (referring to Figure 4.19, extended to nth order $H(s)$):

n inverting scaled integrators (item 2, Table 4.2) for the n SFG branches with weight $-1/s$
1 inverting amplifier at the input end to provide $-V_i$
1 inverting amplifier at the output end to make available both V_o and $-V_o$

The number of op amps can be reduced if the restriction of using single-ended op amp is removed. Table 4.3 describes several differential-input op amp circuits suitable for use in the state variable active filters.

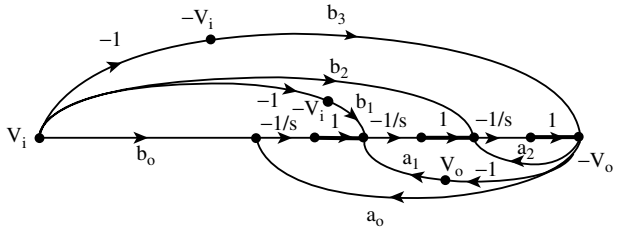
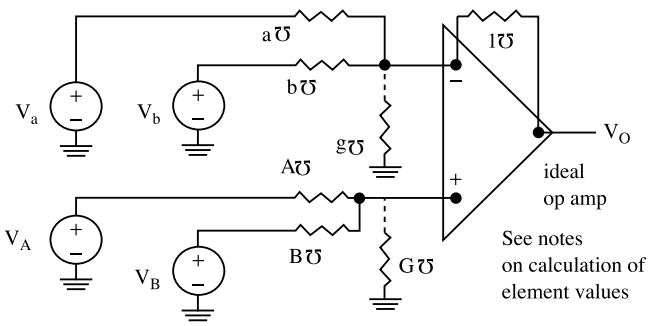


FIGURE 4.19 A modification of Figure 4.16(b) to use all positive *a*'s and *b*'s.

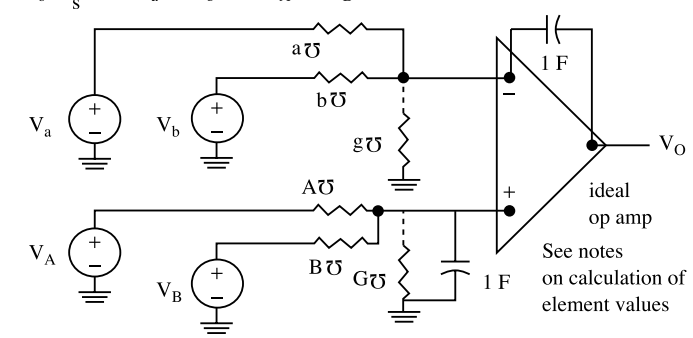
TABLE 4.3 Differential-Input Op Amp Circuit

(1) Bi-polarity scaled summer $V_o = -(aV_a + bV_b) + (AV_A + BV_B)$



(2) Bi-polarity scaled summing integrator

$$V_o = \frac{1}{s}[-(aV_a + bV_b) + (AV_A + BV_B)]$$



Note: Calculation of element values in Table 4.3[7].

(i) The initial design uses 1Ω resistance or 1 F capacitance as the feedback element.

(ii) Either the g mho conductance or the G mho conductance (not both) is connected.

Choose the values of g or G such that the sum of all conductances connected to the inverting input terminal equals the sum of all conductances connected to the noninverting input terminal.

(iii) Starting with the initial design, one may magnitude-scale all elements connected to the inverting input terminal by one factor, and all elements connected to the noninverting input terminal by the same or a different factor.

If differential-input op amps are used, then the number of op amps required for the realization of Eq. (4.1) (with $m = n$) is reduced to $(n + 1)$ for Figure 4.16(a) and n for Figure 4.16(b). The breakdowns are as follows:

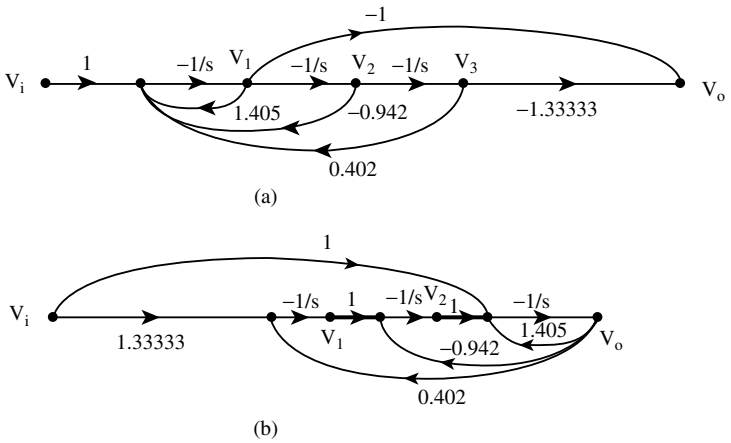


FIGURE 4.20 Two SFGs realizing the transfer function of Eq. (4.7).

For the controllable canonical form SFG of Figure 4.16(a):

$n-1$ inverting integrators (item 2, Table 4.2 with one input) for the n SFG branches with weight $-1/s$, except the leftmost

1 bipolarity-scaled summing integrator (item 2, Table 4.3) for the leftmost SFG branch with weight $-1/s$

1 bipolarity-scaled summer (item 1, Table 4.3) to obtain V_o

For the observable canonical form SFG of Figure 4.16(b):

n bipolarity scaled summing integrator (item 2, Table 4.3), one for each SFG branch with weight $-1/s$

To construct the op amp circuit, one refers to the SFG of Figure 4.14 and obtains the expression relating the output of each op amp to the outputs of other op amps. After that is done, refer to Table 4.3, pick the appropriate component circuits, and connect them as specified by the SFG. The next example outlines the procedure of utilizing differential-input type op amps to reduce the total number of op amps to $(n+1)$ or n.

Example 6. Design a state-variable active low-pass filter to meet the following requirements: magnitude response is of the inverse Chebyshev type

$$\alpha_{\max} = 0.5 \text{ dB}, \qquad \alpha_{\min} = 20 \text{ dB}, \qquad \alpha(\omega_s) = \alpha_{\min}$$

$$\omega_p = 333.33 \text{ rad/s}, \qquad \omega_s = 1000 \text{ rad/s}$$

Solution. Using the method described in Chapter 71, the *normalized* transfer function (i.e., $\omega_s = 1$ rad/s) is found to be

$$H(s) = \frac{V_o}{V_i} = \frac{K\left(s^2 + 1.33333\right)}{s^3 + 1.40534s^2 + 0.94200s + 0.40196} \tag{4.7}$$

The SFGs for this $H(s)$ are simply obtained from Figure 4.16 by removing the two branches having weights b_3 and b_1.The results are shown in Figure 4.20(a) for the case $K = 1$, and in Figure 4.20(b) for the case $K = -1$. A four-op-amp circuit for the normalized $H(s)$ may be constructed in accordance with the SFG of Figure 4.20(a). The component op amp circuits are selected from Table 4.2 and 4.3 in the following manner:

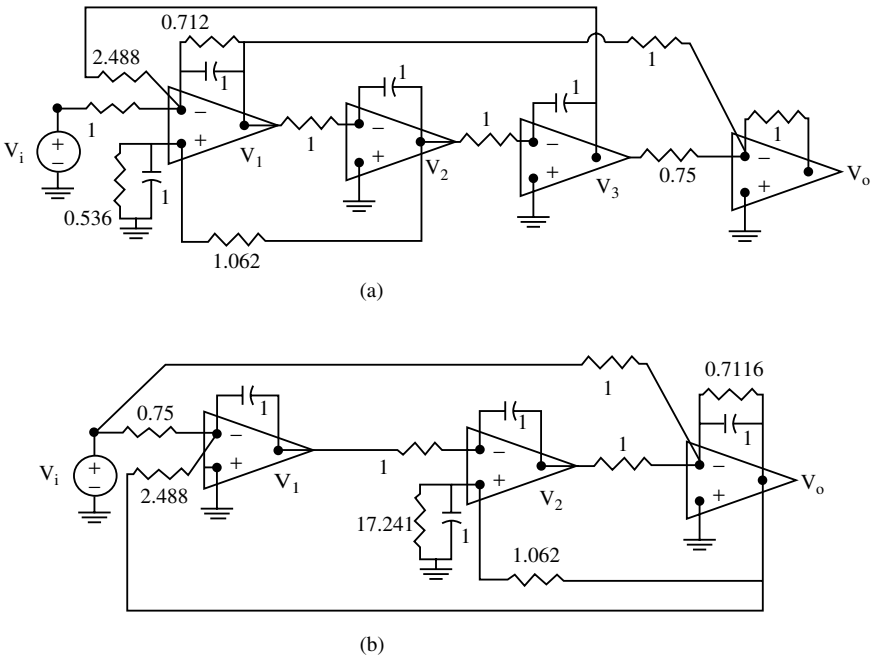


FIGURE 4.21 Two realizations of the third-order, inverse Chebyshev low-pass filter of Example 6. Element values are in kΩ and μF.

Relationship from SFG	Component op amp circuit
$V_1 = -\frac{1}{s}\left(V_i + 1.405V_1 + 0.402V_3 - 0.942V_2\right)$	item 2, Table 4.3
$V_2 = -\frac{1}{s}V_1$	item 2, Table 4.2
$V_3 = -\frac{1}{s}V_2$	item 2, Table 4.2
$V_o = -V_1 - 1.33333\,V_3$	item 1, Table 4.2

After connecting these four-component op amp circuits as described in Figure 4.20(a), and frequency-scaling the whole circuit by 1000, and magnitude-scaling by 1000, the final op amp circuit meeting the low-pass filter specifications is shown in Figure 4.21(a).

In a similar manner, a three-op-amp circuit for the normalized $H(s)$ may be constructed in accordance with the SFG of Figure 4.20(b). The final op amp circuit meeting the lowpass filter specifications is shown in Figure 4.21(b). Both circuits in Figure 4.21 achieve a gain constant $|K| = 1$ in Eq. (4.7). Should a different value of $|K| = 1/\alpha$ be desired, it is only necessary to multiply the values of all resistors connected to the input V_i by α.

When the method of this subsection is applied to a second order transfer function, the resultant op amp circuit is called a *state variable biquad*. Biquads and first order op amp circuits are used as the basic building blocks in the synthesis of a general nth order transfer function by the "cascade" approach. Depending on the SFGs chosen and the types of op amps allowed (single-ended or differential-input), a state variable biquad may require from 2 to 5 op amps. Some special but useful state variable biquads are listed in Table 4.4 for reference purposes.

TABLE 4.4 Some Special State-Variable Biquads

Normalized transfer functions

(1) Lowpass

$$\frac{V_{LP}}{V_i} = \frac{1}{s^2 + \frac{1}{Q}s + 1}$$

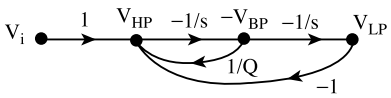
(2) Bandpass

$$\frac{V_{BP}}{V_i} = \frac{s}{s^2 + \frac{1}{Q}s + 1}$$

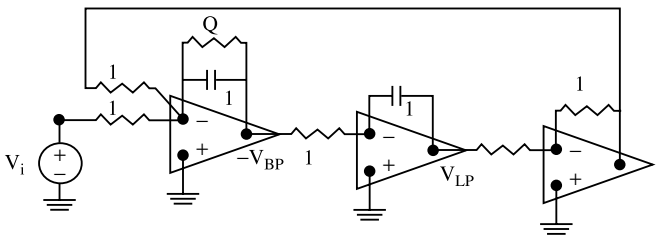
(3) Highpass

$$\frac{V_{HP}}{V_i} = \frac{s^2}{s^2 + \frac{1}{Q}s + 1}$$

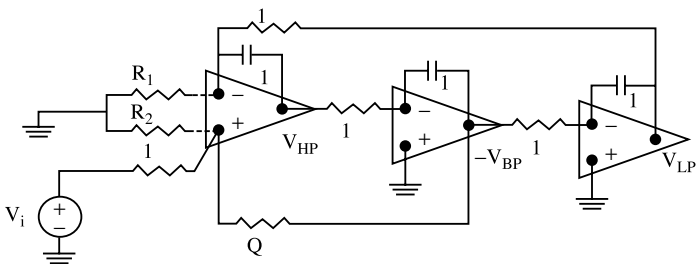
Signal flow graph for transfer functions (1) - (3)



Op amp circuits for (1) – (2). Available outputs: LP and BP



Op amp circuits for (1) – (3). Available outputs: LP, BP and HP



Either R1 or R2 is connected. See notes in Table 20.3 for the determination of their values.

All the SFGs used in the previous examples are of the two types (controllable and observable canonical forms) illustrated in Figure 4.16; however, many other possible SFGs produce the same transfer function. For example, a third-order, low-pass Butterworth or Chebyshev filter has an all-pole transfer function.

$$H(s) = \frac{V_o}{V_i} = \frac{K}{s^3 + a_2 s^2 + a_1 s + a_o} \tag{4.8}$$

A total of six SFGs may be constructed in accordance with Eq. (4.2) to produce the desired $H(s)$. These are illustrated in Figure 4.22. Among these, six SFGs only two have been chosen for consideration in this section.

Similarly, for a fourth-order, low-pass Butterworth or Chebyshev filter, a total of 20 SFGs may be constructed. The reader should consult References [8–9] for details.

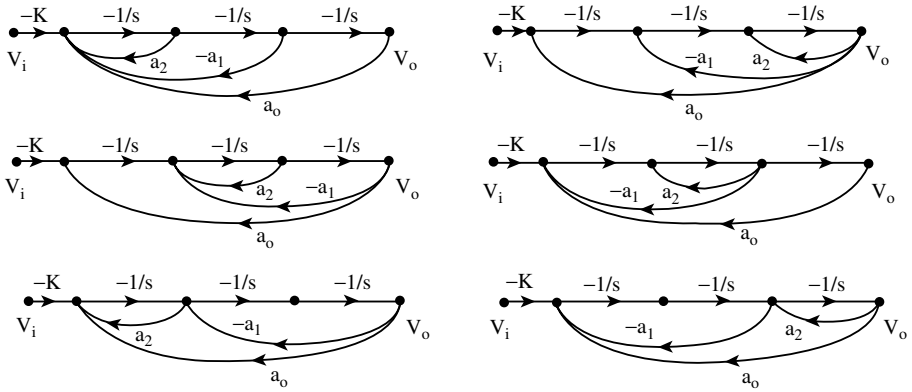


FIGURE 4.22 Six SFGs realizing a third-order, all-pole transfer function.

References

[1] G. Daryanani, *Principles of Active Network Synthesis and Design*, New York: John Wiley & Sons, 1976.

[2] G. C. Temes and J. W. LaPatra, *Introduction to Circuit Synthesis and Design*, New York: McGraw-Hill, 1977.

[3] M. E. Van Valkenburg, *Analog Filter Design*, New York: Holt, Rinehart & Winston, 1982.

[4] W. K. Chen, *Passive and Active Filters*, New York: John Wiley & Sons, 1986.

[5] L. P. Huelsman, *Active and Passive Analog Filter Design*, New York: McGraw-Hill, 1993.

[6] W. J. Kerwin, L. P. Huelsman, and R. W. Newcomb, "State variable synthesis for insensitive integrated circuit transfer functions," *IEEE J. Solid Circuits*, vol. SC-2, pp. 87–92, Sep. 1967.

[7] P. M. Lin, "Simple design procedure for a general summer," *Electron. Eng.*, vol. 57, no. 708, pp. 37–38, Dec. 1985.

[8] N. Fliege, "A new class of canonical realizations for analog and digital circuits," *IEEE Proc. 1984 International Symposium on Circuits and Systems*, pp. 405–408, May 1984.

[9] P. M. Lin, "Topological realization of transfer functions in canonical forms," *IEEE Trans. Automatic Control*, vol. AC-30, pp. 1104–1106, Nov. 1985.

5

Analysis in the Frequency Domain

Jiri Vlach
University of Waterloo, Canada

John Choma, Jr.
University of Southern California

5.1 Network Functions

Jiri Vlach

Definition of Network Functions

Network functions can be defined if the following constraints are satisfied:

1. The network is linear.
2. It is analyzed in the frequency domain using the Laplace transform.
3. All initial voltages and currents are zero (zero state conditions).

This chapter demonstrates how the various functions can be derived, but first we introduce some explanations and definitions. If we analyze any linear network, we can take as output any nodal voltage, or a difference of any two nodal voltages; denote such as output voltage by V_{out}. We can also take as the output a current through any element of the network; we call it output current, I_{out}. If the network is excited by a voltage source, E, then we can also calculate the current delivered into the network by this source; this is the input current, I_{in}. If the network is excited by a current source, J, then the voltage across the current source is the input voltage, V_{in}.

Suppose that we analyze the network and keep the letter E or J in our derivations. Then, we can define the following network functions:

$$\begin{aligned} &\text{Voltage transfer function,} && T_v = \frac{V_{out}}{E} \\ &\text{Input admittance,} && Y_{in} = \frac{I_{in}}{E} \\ &\text{Transfer admittance,} && Y_{tr} = \frac{I_{out}}{E} \end{aligned} \tag{5.1}$$

Current transfer function, $T_i = \dfrac{I_{out}}{J}$

Input impedance, $Z_{in} = \dfrac{V_{in}}{J}$

Transfer impedance, $Z_{tr} = \dfrac{V_{out}}{J}$

Output impedance or output admittance are also used, but the concept is equivalent to the input impedance or admittance. The only difference is that, for calculations, the source is placed temporarily at a point from which the output normally will be taken. In the Laplace transform, it is common to use capital letters, *V* for voltages and *I* for currents. We also deal with impedances, *Z*, and admittances, *Y*. Their relationships are

$$V = ZI \qquad I = YV$$

The impedance of a capacitor is $Z_C = 1/sC$, the impedance of an inductor is $Z_L = sL$, and the impedance of a resistor is R. The inverse of these values are admittances: $Y_C = sC$, $Y_L = 1/sL$, and the admittance of a resistor is $G = 1/R$.

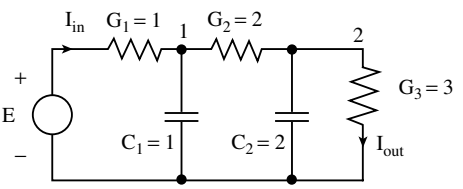


FIGURE 5.1

To demonstrate the derivations of the above functions two examples are used. Consider the network in Figure 5.1, with input delivered by the voltage source, *E*. By Kirchhoff's current law (KCL), the sum of currents flowing *away* from node 1 must be zero:

$$(G_1 + sC_1 + G_2)V_1 - G_2V_2 - EG_1 = 0$$

Similarly, the sum of currents flowing away from node 2 is

$$-V_1G_2 + (G_2 + sC_2 + G_3)V_2 = 0$$

The independent source is denoted by the letter *E*, and is assumed to be known. In mathematics, we transfer known quantities to the right. Doing so and collecting the equations into one matrix equation results in

$$\begin{bmatrix} G_1 + G_2 + sC_1 & -G_2 \\ -G_2 & G_2 + G_3 + sC_2 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} EG_1 \\ 0 \end{bmatrix}$$

If numerical values from the figure are used, this system simplifies to

$$\begin{bmatrix} s+3 & -2 \\ -2 & 2s+5 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \begin{bmatrix} E \\ 0 \end{bmatrix}$$

or

$$\boldsymbol{YV = E}$$

Any method can be used to solve this system, but for the sake of explanation it is advantageous to use Cramer's rule. First, find the determinant of the matrix,

$$D = 2s^2 + 11s + 11$$

To obtain the solution for the variable V_1 (V_2), replace the first (second) column of $\boldsymbol{Y}$ by the right-hand side and calculate the determinant of such a modified matrix. Denoting such a determinant by the letter N with an appropriate subscript, evaluate

$$N_1 = \begin{vmatrix} E & -2 \\ 0 & 2s+5 \end{vmatrix} = (2s+5)E$$

Then,

$$V_1 = \frac{N_1}{D} = \frac{2s+5}{2s^2+11s+11}E$$

Now, divide the equation by E, which results in the voltage transfer function

$$T_v = \frac{V_1}{E} = \frac{2s+5}{2s^2+11s+11}$$

To find the nodal voltage V_2, replace the second column by the elements of the vector on the right-hand side:

$$N_2 = \begin{vmatrix} s+3 & E \\ -2 & 0 \end{vmatrix} = 2E$$

The voltage is

$$V_2 = \frac{N_2}{D} = \frac{2}{2s^2+11s+11}E$$

and another voltage transfer function of the same network is

$$T_v = \frac{V_2}{E} = \frac{2}{2s^2+11s+11}$$

Note that many network functions can be defined for any network. For instance, we may wish to calculate the currents I_{in} or I_{out}, marked in Figure 5.1. Because the voltages are already known, they are used: For the output current $I_{out} = G_3 V_2$ and divided by E

$$Y_{tr} = \frac{I_{out}}{E} = \frac{3V_2}{E} = \frac{6}{2s^2+11s+11}$$

The input current $I_{in} = E - G_1 V_1 = E - V_1 = E(2s^2 + 9s + 6)/(2s^2 + 11s + 11)$ and dividing by E

$$Y_{in} = \frac{I_{in}}{E} = \frac{2s^2+9s+6}{2s^2+11s+11}$$

In order to define the other possible network functions, we must use a current source, J, as in Figure 5.2, where we also take the current through the inductor as an output variable. This method of formulating the network equations is called **modified nodal**. The sum of currents flowing away from node 1 is

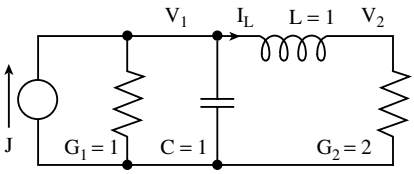


FIGURE 5.2

$$\left(G_1 + sC_1\right)V_1 + I_L - J = 0$$

from node 2 it is

$$G_2 V_2 - I_L = 0$$

and the properties of the inductor are expressed by the additional equation

$$V_1 - V_2 - sLI_L = 0$$

Inserting numerical values and collecting in matrix form:

$$\begin{bmatrix} s+1 & 0 & 1 \\ 0 & 2 & -1 \\ 1 & -1 & -s \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ I_L \end{bmatrix} = \begin{bmatrix} J \\ 0 \\ 0 \end{bmatrix}$$

The determinant of the system is

$$D = -\left(2s^2 + 3s + 3\right)$$

To solve for V_1, we replace the first column by the right-hand side and evaluate the determinant

$$N_1 = \begin{vmatrix} J & 0 & 1 \\ 0 & 2 & -1 \\ 0 & -1 & -s \end{vmatrix} = -\left(2s+1\right)J$$

Then, $V_1 = N_1/D$ and dividing by J we obtain the network function

$$Z_{\text{tr}} = \frac{V_1}{J} = \frac{2s+1}{2s^2+3s+3}$$

To obtain the inductor current, evaluate the determinant of a matrix in which the third column is replaced by the right-hand side: $N_3 = -2J$. Then, $I_L = N_3/D$ and

$$T_i = \frac{I_L}{J} = \frac{2}{s^2+3s+3}$$

In general,

$$F = \frac{\text{Output variable}}{E \text{ or } J} = \frac{\text{Numerator polynomial}}{\text{Denominator polynomial}} \tag{5.2}$$

Any method that may be used to formulate the equations will lead to the same result. One example shows this is true. Reconsider the network in Figure 5.2, but use the admittance of the inductor, $Y_L = 1/sL$, and do not consider the current through the inductor. In such a case, the nodal equations are

$$\left\{1+s+\frac{1}{s}\right\}V_1-\frac{1}{s}V_2=J$$

$$-\frac{1}{s}V_1+\left\{\frac{1}{s}+2\right\}V_2=0$$

We can proceed in two ways:

1. We can multiply each equation by s and thus remove the fractions. This provides the system equation

$$\begin{bmatrix} s^2+s+1 & -1 \\ -1 & 2s+1 \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}=\begin{bmatrix} sJ \\ 0 \end{bmatrix}$$

   The determinant of this matrix is $D = 2s^3 + 3s^2 + 3s$. To calculate V_1, find $N_1 = s(2s + 1)J$. Their ratio is the same as before because one s in the numerator can be canceled against the denominator.
2. If we do not remove the fractions and go ahead with the solution, we have the matrix equation

$$\begin{bmatrix} s+1+1/s & -1/s \\ -1/s & 1/s+2 \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}=\begin{bmatrix} J \\ 0 \end{bmatrix}$$

   The determinant is $D = 2s + 3 + 3/s$ and the numerator for V_1 is $N_1 = (1/s + 2)J$. Taking their ratio

$$V_1=\frac{(1/s+2)J}{2s+3+3/s}=\frac{2s+1}{2s^2+3s+3}J$$

   which is the same result as before.

We conclude that it does not matter which method is used to formulate the equations. The result is always a ratio of two polynomials in the variable s.

Many additional conclusions can be drawn from these examples. The most important result so far is that **all network functions of any given network have the same denominator**. It was easy to discover this property because we used Cramer's rule, with its evaluation by the ratio of two determinants. It should be mentioned at this point that we may have network functions in which some terms of the numerator can cancel against the same terms of the denominator. Such a cancellation represents a mathematical simplification which does not change the validity of the above statement.

Occasionally, the network may have more than one source. In such cases, we apply the superposition principle of linear networks. The contribution to the output can be calculated separately for each source and the results added. All that must be done is to correctly remove those sources which are not considered at the moment. All *unused independent voltage sources* must be replaced by *short circuits*. All *unused independent current sources* are replaced by *open circuits* (removed from the network). Although we did not use dependent sources in our examples, it is necessary to stress that such removal *must not* take place for dependent sources.

Network functions can be used to find responses to any given input signal. First, multiply the network function by E or J; this will give the expression for the output. Afterward, the letter E or J is replaced by the Laplace transform of the signal. For instance, if the signal is a unit step, then the source is replaced by $1/s$. If it is cost ωt, then the source is replaced by the Laplace transform, $s/(s^2 + \omega^2)$, and so on.

In the Laplace transform, one special signal exists, the Dirac impulse, commonly denoted by $\delta(t)$. It can be represented as a rectangular pulse having width T and height $1/T$. The area of the pulse is always 1, even if we go to $\lim T \to 0$, which is the Dirac impulse. Its Laplace transform is 1. Because multiplication by 1 does not change the network function, we conclude that any network function is also the Laplace transform of the network response to the Dirac impulse.

A word of caution: In the network function always divide by the *independent* voltage (current) source. We cannot take two analysis results, for instance V_1 and V_2, derived for Figure 5.1, and take their ratio. This will not be a network function.

Poles and Zeros

Networks with lumped elements have network functions which are always ratios of two polynomials with real coefficients. For some applications the polynomials may be expressed as functions of some (or all) elements, but the principle is unchanged.

Because we have a ratio of two polynomials, the network function can be written in two forms:

$$F = \frac{\sum_{i=o}^{M} a_i s^i}{\sum_{i=o}^{N} b_i s^i} = K \frac{\prod_{i=1}^{M} (s - z_i)}{\prod_{i=1}^{N} (s - p_i)} \tag{5.3}$$

The middle form is what we obtain from analyses similar to those in the examples. Algebraically, a polynomial of order N has exactly N roots. This leads to the form on the right. The multiplicative constant, K, is the ratio

$$K = \frac{a_M}{b_N}$$

and is obtained by dividing each polynomial by the coefficient of its highest power.

It is easy to find roots of a first- and second-order polynomial because formulas are available, but in all other cases iterative methods and a computer are utilized. However, even without actually finding the roots, we can draw a number of important conclusions.

If the highest power of the polynomial is odd, then at least one real root will exist. The other roots may be either real or complex, but if they are complex, then they always appear in complex conjugate pairs. The roots of the numerator are called *zeros*, and those of the denominator are called *poles*. We denote the zeros by

$$z_i = a_i + jb_i$$

where $j = \sqrt{-1}$. Either a or b may be zero. For the poles, we have similarly

$$p_i = c_i + jd_i$$

The polynomial also may have multiple roots. For instance, the polynomial $P(s) = (s + 1)^2 (s + 2)^3$ has a double root at $s = -1$ and a triple root at $s = -2$. The positions of the poles and zeros, with the constant K, completely define the network function and also all network properties. The positions can be plotted in a complex plane, the zeros indicated by small circles and poles by crosses. A multiple pole (zero) is indicated by a number appearing at the cross (circle). Figure 5.3 shows a network function with two complex conjugate zeros on the imaginary axis, two complex conjugate poles, and one double real pole.

As derived previously, all network functions of any given network have the same poles. Their positions depend only on the structure of the network and are independent of the signal or where the signal is applied. Because of this fundamental property, the poles are also called *natural frequencies* of the network.

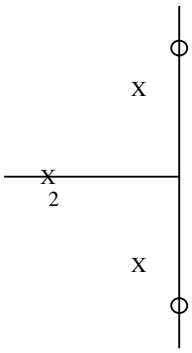


FIGURE 5.3

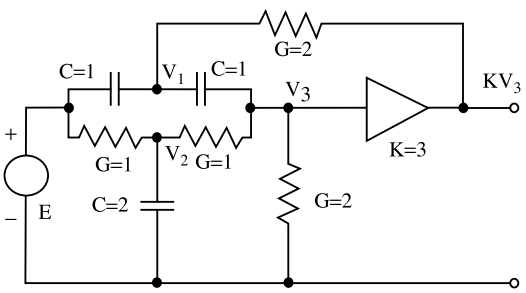


FIGURE 5.4

The zeros depend on the place at which we attach the source and also on the point where we take the output.

It is possible to have networks in which a pole is in exactly the same position as a zero; mathematically, such terms cancel. Figure 5.4 is an example. Writing the sum of currents at nodes 1, 2, and 3, we obtain

$$\begin{aligned}(2s+2)V_1-(s+3)V_3 &= sE\\ (2s+2)V_2-V_3 &= E\\ -sV_1-V_2+(s+3)V_3 &= 0\end{aligned}$$

and in matrix form

$$\begin{bmatrix} 2s+2 & 0 & -(s+3)\\ 0 & 2s+2 & -1\\ -s & -1 & s+3\end{bmatrix}\begin{bmatrix} V_1\\ V_2\\ V_3\end{bmatrix}=\begin{bmatrix} sE\\ E\\ 0\end{bmatrix}$$

By carefully evaluating the determinant we discover that we can keep the term $(2s + 2)$ separate and get $D = (2s + 2)(s^2 + 2s + 5)$. Replacing the third column by the right-hand side, we calculate the numerator $N_3 = (2s + 2)(s^2 + 1)E$. Because the output is KV_3, the voltage transfer function is

$$T_v=\frac{3(2s+2)(s^2+1)}{(2s+2)(s^2+2s+5)}$$

Mathematically, the term $(2s + 2)$ cancels and the network function is sometimes written as

$$T_v = \frac{3\left(s^2+1\right)}{s^2+2s+5}$$

Such cancellation makes the denominator different from other network functions that we might derive for the same network, but it is not a correct way to describe the properties of the network. The cancellation gives the impression that we have a second-order network, while it is actually a third-order network.

Network Stability

Stability of the network depends entirely on the positions of its poles. The following is a list of the conditions in order for the network to be stable, with subsequent explanation of the reasons:

1. The network is stable if all its poles are in the left half of the complex plane.
2. The network is unstable if at least one of its poles is in the right-half plane.
3. The network is marginally stable if all its poles are simple and exactly on the imaginary axis.
4. The network is unstable if it has all poles on the imaginary axis, but at least one of them has multiplicity two or more.

Courses on mathematics teach the process of decomposing a rational function into partial fractions. We show an example with one simple real pole and a pair of simple complex conjugate poles,

$$F(s) = \frac{3s^2+8s+6}{(s+1)\left(s^2+2s+2\right)} = \frac{1}{s+1} + \frac{1+j}{s+1+j} + \frac{1-j}{s+1-j}$$

The poles are $p_1 = -1$ and $p_{2,3} = -1 \pm j$, all with negative real parts and all lying in the left-half plane. Partial fraction decomposition is on the right of the preceding equation. It is always true, for any lumped network, that the decomposition for a real pole has a real constant in the numerator. Complex poles always appear in complex conjugate pairs and the decomposition constants, if complex, also are complex conjugate. Once such a decomposition is available, tables can be used to invert the functions into time domain. The decomposition may be quite a laborious process, however, only a few types of terms need be considered for lumped networks. All are collected in Table 5.1. Each time domain expression is multiplied by unit step, $u(t)$, which is zero for $t < 0$ and is one for $t \geq 0$. Such multiplication correctly expresses the fact that the time functions start at $t = 0$.

Formula one in Table 5.1 shows that a real, single pole in the left-half plane will lead to a time-domain function which decreases as e^{-ct}. This response is called stable. If $c = 0$, then the response becomes $u(t)$.

TABLE 5.1

Formula	Laplace Domain	Time Domain
1	$\frac{K}{s+c}$	$Ke^{-ct}u(t)$
2	$\frac{K}{(s+c)^n}$	$K\frac{t^{n-1}}{(n-1)!}e^{-ct}u(t)$
3	$\frac{A+jB}{s+c+jd} + \frac{A-jB}{s+c-jd}$	$2e^{-ct}(A\cos dt + B\sin dt)u(t)$
4	$\frac{A+jB}{(s+c+jd)^n} + \frac{A-jB}{(s+c-jd)^n}$	$\frac{2t^{n-1}}{(n-1)!}e^{-ct}(A\cos dt + B\sin dt)u(t)$

Should the pole be in the right-half plane, the exponent will be positive and e^{ct} will grow rapidly and without bound. This network is said to be unstable.

Formula two shows what happens if the pole is real, with multiplicity n. If it is in the left-half plane, then t^{n-1} is a growing function, but e^{-ct} decreases faster, and for large t the result tends to zero. The function is still stable.

Formula three considers the case of two simple complex conjugate poles. Their real parts influence the exponent, and the imaginary parts contribute to oscillations. If the real part is negative, the oscillations will be damped, the response will become zero for large t, and the network will be stable. If the real part is zero, then the oscillations continue indefinitely with constant amplitude. For the positive real part, the network becomes unstable.

Formula four considers a pair of multiple complex conjugate poles. As long as the real part is negative, the oscillations will decrease with time and the network will be stable. If a real part is zero or positive, the network is unstable because the oscillations will grow.

Initial and Final Value Theorems

Finding the poles and evaluating the time domain response is a complicated process, which normally requires the use of a computer. It is, therefore, advisable to use all possible steps that may provide information about the network behavior without actually finding the full time-domain response.

Two Laplace transform theorems help in finding how the network behaves at $t = 0$ and at $t \to \infty$. Both theorems are derived from the Laplace transform formula for differentiation,

$$\int_{0^-}^{\infty} f'(t)e^{-st}dt = sF(s) - f(0^-) \tag{5.4}$$

where 0^- indicates that we are considering the instant just before the signal is applied. If we let $s \to 0$, then $e^0 = 1$, and the integral of the derivative becomes the function itself. Inserting the integrating limits we get

$$f(\infty) - f(0^-) = \lim_{s\to 0}\left[sF(s) - f(0^-)\right]$$

Cancelling $f(0^-)$ on both sides, we arrive at the *final value theorem*

$$\lim_{t\to\infty} f(t) = \lim_{s\to 0} sF(s) \tag{5.5}$$

Another possibility is to let $s \to \infty$; then e^{-st} in (5.4) will be zero and the whole left side becomes zero. This can be written as

$$0 = \lim_{s\to\infty}\left[sF(s) - f(0^-)\right]$$

and because $f(0^-)$ is nothing but the limit of $f(t)$ for $t \to 0^-$, we obtain the *initial value theorem*

$$\lim_{t\to 0} f(t) = \lim_{s\to\infty} sF(s) \tag{5.6}$$

Note the similarity of the two theorems; we will apply them to the function used in the previous section. Consider

$$sF(s) = \frac{3s^3 + 8s^2 + 6s}{s^3 + 3s^2 + 4s + 2}$$

TABLE 5.2

Laplace Domain $D = s^3 + 3s^2 + 4s + 2$	Time Domain	$s \to \infty$ $t = 0$	$s = 0$ $t \to \infty$
$F(s) = 1/D$	$f(t) = e^{-t}(1 - \cos t)u(t)$	0	0
$G(s) = sF(s) = s/D$	$g(t) = f'(t) = e^{-t}(-1 + \cos t + \sin t)u(t)$	0	0
$H(s) = s^2F(s) = s^2/D$	$h(t) = f''(t) = e^{-t}(1 - 2\sin t)u(t)$	1	0
$K(s) = s^3/D$	$k(t) = f'''(t) = \delta(t) + e^{-t}(2\sin t - 2\cos t - 1)u(t)$	$\delta(t)$	0

If we take any large value of s, the highest powers will dominate and in limit, for $s \to \infty$, we get 3. This is the value of the time-domain response at $t = 0$. The limit for $s = 0$ is zero, and from the final value theorem we know that $f(t)$ will be zero for $t \to \infty$.

To extract still more information, use the example collected in Table 5.2. Scrolling down the table, each Laplace domain function is s times that above it. Each multiplication by s means differentiation in the time domain, as follows from (5.4). Scrolling down the second column of Table 5.2, each function is the derivative of that above it. To apply the limiting theorems, take the Laplace domain formula, which is one level lower, and insert the limits. The limiting is also shown and is confirmed by inserting either $t = 0$ or $t \to \infty$ into the time functions.

Although the two theorems are useful, the final value theorem is valid *only if the function is stable*. Consider the unstable function with two poles in the right-half plane

$$F_1(s) = \frac{1}{(s+1)(s-1+j)(s-1-j)} = \frac{1}{s^3 - s^2 + 2}$$

Its time-domain response is

$$f_1(t) = \frac{1}{5}\left[e^{-t} + e^{+t}\left(2\sin^t - \cos^t\right)\right]u(t)$$

and the term e^{+t} will cause the function to grow for large t. If the final value theorem is applied, we consider

$$sF_1(s) = \frac{s}{s^3 + s^2 + 2}$$

Inserting $s = 0$, the theorem predicts that the time function will approach zero for large t. This is disappointing, but some additional simple rules can be applied. The function is unstable if some coefficients of the denominator are missing, or if the denominator coefficients do not all have the same sign (all + or all –). Such situations are easily detected, but if all coefficients have the same sign, nothing can be said about stability. Additional theorems exist (e.g., Hurwitz theorem), but if in doubt, it is probably simplest to go to the computer and find the poles.

5.2 Advanced Network Analysis Concepts

John Choma, Jr.

Introduction

The systematic analysis of an electrical or electronic network entails formulating and solving the relevant Kirchhoff equations of equilibrium. The analysis is conducted to acquire a theoretically sound understanding of circuit responses. Such an understanding minimally delineates the dynamical effects of

topology, controllable circuit branch variables, and observable parameters for active devices embedded in the circuit. It also illuminates circuit node and branch impedances to which the relevant responses of the circuit undergoing investigation are especially sensitive. Unfortunately, the complexity of modern networks, and particularly integrated analog electronic circuits, often inhibits the mathematical tractability that underpins an engineering understanding of circuit behavior. It is therefore not surprising that when mathematical analyses accompany a computer-assisted circuit design venture, the subcircuits identified for manual study are simplified representations of the corresponding subcircuits in the draft design solution. Unless care is exercised, these approximations can mask a satisfying understanding, and they can even lead to erroneous results.

Analytical and modeling approximations notwithstanding, the key to assimilating a satisfying understanding of the electrical characteristics of complex circuits is appropriate studies of simpler partitions of these circuits. To this end, Kron [1, 2] has provided, and others have explained and reinforced [3–5], an elegant theory that allows the circuit response solutions of these network partitions to be coalesced so that the desired response of the interconnected circuit is reconstructed exactly. Aside from formalizing an analytical mechanism for studying complicated circuits in terms of the solutions gleaned for more manageable subcircuits of the composite network [6], Kron's work allows for a computationally efficient study of feedback network responses. The theory also allows for the investigation of the sensitivity of overall network performance with respect to both small and large parametric changes [7]. In view of the exclusive focus on linear circuits in this section, it is worth interjecting that a form of Kron's partitioning theory is also applicable to certain classes of nonlinear circuits [8].

Fundamental Network Analysis Concepts

The derivation of Kron's formula, as well as the development of a general methodology for applying Kron's partitioning mechanism to the analyses of complex circuits, requires a fundamental understanding of the classical techniques exploited in the analysis of linear networks. Such an understanding begins by considering the $(n + 1)$ node, b branch linear network abstracted in Figure 5.5(a). The input port, which is defined by the node pair, 1-2, is excited by a signal source whose Thévenin voltage is V_s and whose Thévenin impedance is Z_s. In response to this excitation, a load voltage, V_L, is developed across a load impedance, Z_L, which shunts the output port consisting of the node pair, 3-4. Two other nodes, nodes m and p, are explicitly delineated for future reference. In response to the applied signal source, the voltage across the input port is V_I, while the voltage established across the node pair, m-p is V_k. In addition, the reference, or ground, node is labeled node 0. Either node 2, node 4, or both of these nodes can be incident with the ground node; that is, the signal source and/or the load impedance can be terminated to the

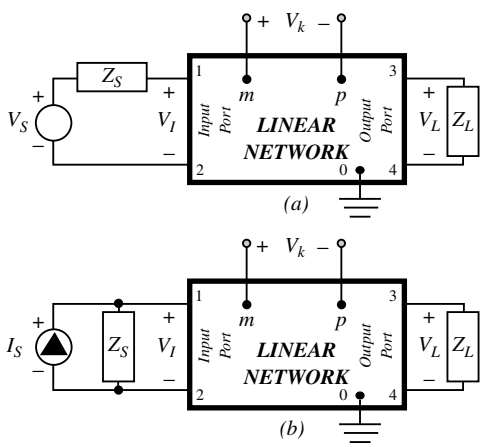


FIGURE 5.5 (a) Generalized linear network driven by a voltage source. (b) The network of (a), but with the signal excitation represented by its Norton equivalent circuit.

network ground. The diagram in Figure 5.5(b) is identical to that of Figure 5.5(a) except for the fact that the applied signal source is represented by its Norton equivalent circuit, where the Norton signal current, I_s, is

$$I_S = \frac{V_S}{Z_S} \tag{5.7}$$

Assuming that a nodal admittance matrix exists for the linear $(n + 1)$ node network at hand, the n equilibrium KCL equations can be expressed as the matrix relationship

$$\mathbf{J} = \mathbf{YE} \tag{5.8}$$

where $\mathbf{J}$ is an n-vector whose ith entry, J_i, is an independent current flowing into the ith circuit node, $\mathbf{E}$ is an n-vector of node voltages such that its ith entry, E_i, is ith node voltage referenced to network ground, and $\mathbf{Y}$, a square matrix of order n, is the nodal admittance matrix of the circuit. If $\mathbf{Y}$ is nonsingular, the node voltages follow as

$$\mathbf{E} = \mathbf{Y}^{-1}\mathbf{J} \tag{5.9}$$

Note that (5.9) is useful symbolically, but not necessarily computationally. In particular, (5.9) shows that the n node voltages of the $(n + 1)$ node, b branch network of Figure 5.5 can be straightforwardly computed in terms of the known independent current source vector and the parameters embedded in the network nodal admittance matrix. In an actual analytical environment, however, the nodal admittance matrix is rarely formulated and inverted. Instead, some or all of the n node voltages of interest are determined merely by algebraically manipulating and solving either the n independent KCL equations or the $(b - n)$ independent Kirchhoff's voltage law (KVL) equations that are required to establish the equilibrium conditions of the subject network.

If the n vector, $\mathbf{E}$, is indeed evaluated, all n independent node voltages are known, because

$$\mathbf{E}^{\mathrm{T}} = \left[E_1, E_2, E_3, \ldots, E_m, \ldots, E_p, \ldots, E_n\right] \tag{5.10}$$

where the superscript T indicates the operation of *matrix transposition*. In general, E_i, for $i = 1, 2, \ldots, n$, is the voltage developed at node i with respect to ground. It follows that the voltage between any two nodes derives directly from the network solution inferred by (5.9). For example, the input port voltage, V_I, is $(E_1 - E_2)$, the output port voltage, V_L, is $(E_3 - E_4)$, and the voltage, V_k, from node m to node p is $V_k = (E_m - E_p)$.

The calculation of the voltage appearing between any two circuit nodes can be formalized with the help of the generalized network diagrammed in Figure 5.6 and through the introduction of the *connection vector* concept. In particular, let $\mathbf{A}_{ij}$ denotes the $(n \times 1)$ connection vector for the port defined by the node pair, i-j. Moreover, let the voltage, V, at node i be taken as positive with respect to node j, and allow a current, I (which may be zero), to flow into node i and out of node j, as indicated in the diagram. Then, the elements of the connection vector, $\mathbf{A}_{ij}$, are all zero except for a + 1 in its ith row and a – 1 in

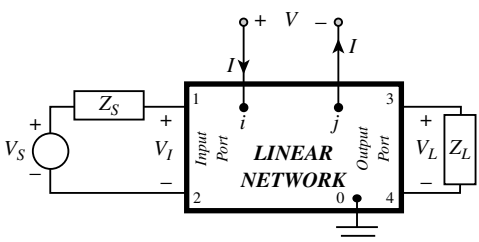


FIGURE 5.6 Generalized network diagram used to define the connection vector concept.

its *j*th row. If node *j* is the reference node, all elements of $\mathbf{A}_{ij}$, which in the case can be written simply as $\mathbf{A}_i$, are zero except for the *i*th row element, which remains +1. Thus, $\mathbf{A}_{ij}$ has the form

$$\mathbf{A}_{ij}^{\mathrm{T}} = \begin{bmatrix} 0 & 0 & \cdots & \underset{\substack{\uparrow \\ i\text{th column}}}{+1} & \cdots & \underset{\substack{\uparrow \\ j\text{th column}}}{-1} & \cdots & \underset{\substack{\uparrow \\ n\text{th column}}}{0} \end{bmatrix} \tag{5.11}$$

For the special case in which a circuit branch element interconnects every pair of circuit nodes, $\mathbf{A}_{ij}$ is the appropriate column of the *node to branch incidence matrix*, which is a rectangular matrix of order $(n \times b)$, for the $(n + 1)$ node, b branch network at hand [9].

Returning to the calculation of V_I, V_L, and V_k, it follows from (5.9) through (5.11) that

$$V_I = E_1 - E_2 = \mathbf{A}_{12}^{\mathrm{T}}\mathbf{E} = \mathbf{A}_{12}^{\mathrm{T}}\mathbf{Y}^{-1}\mathbf{J} \tag{5.12}$$

$$V_L = E_3 - E_4 = \mathbf{A}_{34}^{\mathrm{T}}\mathbf{E} = \mathbf{A}_{34}^{\mathrm{T}}\mathbf{Y}^{-1}\mathbf{J} \tag{5.13}$$

and

$$V_k = E_m - E_p = \mathbf{A}_{mp}^{\mathrm{T}}\mathbf{E} = \mathbf{A}_{mp}^{\mathrm{T}}\mathbf{Y}^{1}\mathbf{J} \tag{5.14}$$

Assuming that I_s is the only independent source of excitation in the network of Figure 5.5

$$\mathbf{J} = \mathrm{A}_{12} I_S \tag{5.15}$$

which is the mathematical equivalent of the observation that the Norton source current, I_S, is entering node 1 of the network and leaving node 2. Accordingly,

$$V_I = \left(\mathbf{A}_{12}^{\mathrm{T}}\mathbf{Y}^{-1}\mathbf{A}_{12}\right) I_S \tag{5.16}$$

$$V_L = \left(\mathbf{A}_{34}^{\mathrm{T}}\mathbf{Y}^{-1}\mathbf{A}_{12}\right) I_S \tag{5.17}$$

and

$$V_k = \left(\mathbf{A}_{mp}^{\mathrm{T}}\mathbf{Y}^{-1}\mathbf{A}_{12}\right) I_S \tag{5.18}$$

Several noteworthy features are implicit to the foregoing three relationships. First, each of the three parenthesized matrix products on the right-hand sides of the equations is a scalar. This observation follows from the facts that a transposed connection vector is a row matrix of order $(1 \times n)$, the inverse nodal admittance matrix is an n-square, and a connection vector is an n-vector. Second, these scalar products represent transimpedances from the input port to the port at which the voltage of interest is extracted. In the case of (5.16), the ratio, $V_1 I_S$, is actually the impedance, Z_{SS}, seen by the Norton current, I_S; that is,

$$Z_{SS} \triangleq \frac{V_S}{I_S} = \left(\mathbf{A}_{12}^{\mathrm{T}}\mathbf{Y}^{-1}\mathbf{A}_{12}\right) \tag{5.19}$$

where, as asserted previously, I_S is presumed to be the only source of energy applied to the network undergoing study. Similarly,

$$Z_{LS} \triangleq \frac{V_L}{I_S} = \left(\mathbf{A}_{34}^{\mathrm{T}}\mathbf{Y}^{-1}\mathbf{A}_{12}\right) \tag{5.20}$$

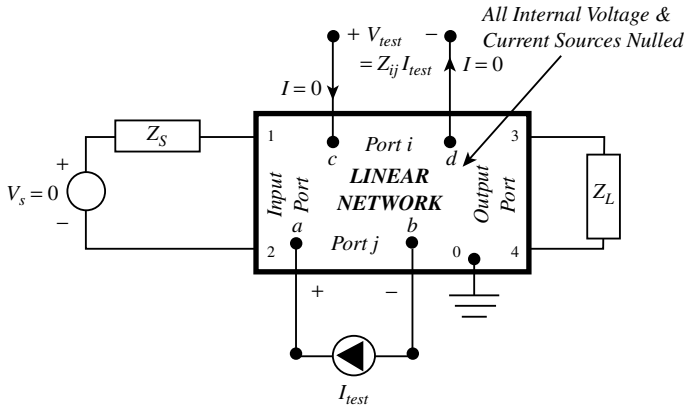


FIGURE 5.7 An illustration of a practical manual technique for computing the transimpedance between any port j to any port i of a linear network.

is the transimpedance from the input port to the output port, while

$$Z_{kS} \triangleq \frac{V_k}{I_S} = \left(\mathbf{A}_{mp}^{\mathrm{T}} \mathbf{Y}^{-1} \mathbf{A}_{12}\right) \tag{5.21}$$

is the transimpedance from the input port to the port defined by the node pair, *m-p*.

The impedance in (5.19) and the transimpedances given by (5.20) and (5.21) are cast as explicit algebraic functions of the inverse of the network nodal admittance matrix. However, similar to the node voltages in (5.19) and (5.10), network transimpedances are rarely calculated manually through an actual delineation and inversion of the nodal admittance matrix. Instead, they usually derive from a straightforward analysis of the considered network, subject to the proviso that all excitations applied to the subject network, save for the single test current source, are reduced to zero. For example, in the abstraction shown in Figure 5.7, the transimpedance, Z_{ij}, from any port j to any port i is

$$Z_{ij} \triangleq \left.\frac{V_{\text{test}}}{I_{\text{test}}}\right|_{\text{all independent sources=0}} \tag{5.22}$$

For the case of $j = i$, this transimpedance becomes the effective impedance seen at port i by the test current source. In view of the preceding discussion, and the node pairs indicated in Figure 5.7, the transimpedance (or impedance) quantity that derives from (5.22) is identical to the matrix relationship

$$Z_{ij} = \left(\mathbf{A}_{cd}^{\mathrm{T}} \mathbf{Y}^{-1} \mathbf{A}_{ab}\right) \tag{5.23}$$

The last result highlights the fact that all network transimpedances are directly related to the inverse of the nodal admittance matrix. Hence, these transimpedances are inversely proportional to the determinant, $\Delta Y(s)$, of the admittance matrix, **Y**. It follows that the poles of all transimpedances and effective port impedances are the roots of the characteristic polynomial

$$\det(\mathbf{Y}) \triangleq \Delta Y(s) = 0 \tag{5.24}$$

Note from (5.7), (5.17), and (5.20) that the voltage gain of the considered linear network is

$$\frac{V_L}{V_S} = \frac{Z_{LS}}{Z_S} \tag{5.25}$$

Thus, if the source impedance in Figure 5.5 is a real number, $Z_S = R_S$, the roots of (5.24) also comprise the poles of the voltage transfer function and, indeed, of the linear network.

Kron's Formula

Assume now that the network depicted in Figure 5.5 has been analyzed in the sense that all network node voltages developed in response to the signal source have been determined. Assume further that subsequent to this analysis, an impedance, Z_k, appended to nodes m and p, as shown in Figure 5.8. In addition to causing a current, I, to flow into node m and out of node p, this additional branch element is likely to perturb the values of all of the originally computed circuit node and circuit branch voltages. The matrix, $\mathbf{E}'$, of new node voltages can be evaluated for the modified topology in Figure 5.8 by determining the new nodal admittance matrix $\mathbf{Y}'$, and the reapplying (5.9). The tedium associated with a second network analysis, along with the inefficiency of discarding the results of a study performed on a network whose topology differs only modestly from that of the original configuration, can be circumvented through the use of *Kron's theorem*. As illuminated next, this theorem derives from a methodical application of such classical concepts as the theories of superposition, substitution, and Thévenin. In addition to providing a computationally efficient mechanism for determining $\mathbf{E}'$, Kron's technique allows for a direct comparison of $\mathbf{E}'$ to the matrix, $\mathbf{E}$, of original node voltages. It therefore allows for a convenient response sensitivity analysis with respect to the appended branch element.

It is appropriate to interject that the problem postulated previously possesses more than mere academic interest. It is, in fact, a problem that is commonly encountered, for example, in the analysis of electronic circuits. In order to linearize these circuits around specified quiescent operating points, it is necessary to supplant the utilized active devices by small signal equivalent circuits. Such models are invariably simplified, often through the tacit neglect of presumably noncritical branch elements, to mitigate analytical complexity and tedium. Thus, while the circuit properly identified for investigation might be of the topological form appearing in Figure 5.8, the circuit actually subjected to manual circuit analysis is likely the reduced structure depicted in Figure 5.5; that is, the ostensibly noncritical impedance, Z_k is removed in the interest of analytical tractability. Questions naturally arise in regard to the degree of error incurred by the invoked circuit simplification. Kron's method, as developed next, answers these questions in terms of the results already deduced for the approximate network and without requiring explicit analytic results for the "exact" network.

The process of evaluating the perturbation on network node voltages incurred by the action of shunting nodes m and p in the circuit of Figure 5.5 by the impedance Z_k begins by determining the Thévenin equivalent circuit that drives the appended branch. To this end, Z_k is removed in the diagram of Figure 5.8, thereby collapsing the network to Figure 5.5(a). The relevant Thévenin voltage, V_{th}, at the node pair, m-p, is, in fact, V_k, as defined by (5.18). Recalling (5.21), this voltage is

$$V_{\text{th}} \equiv V_k = \left(\mathbf{A}_{mp}^{\text{T}}\mathbf{Y}^{-1}\mathbf{A}_{12}\right)I_S = Z_{kS}I_S \tag{5.26}$$

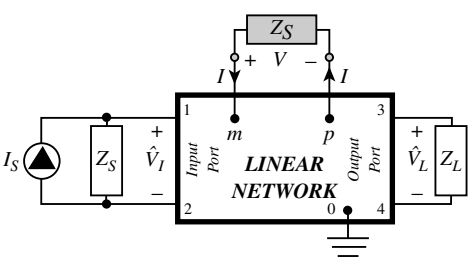


FIGURE 5.8 The inclusion of an impedance, Z_k, between nodes m And p, subsequent to the analysis of the network in Figure 5.5.

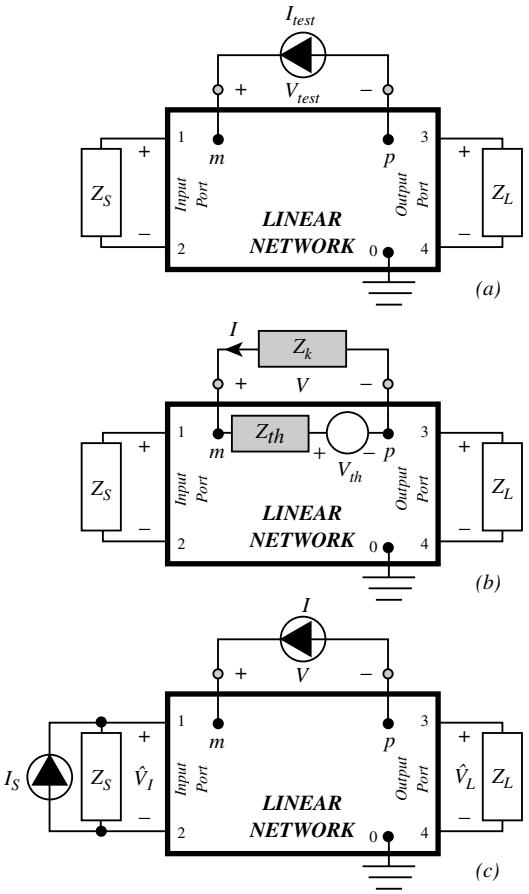


FIGURE 5.9 (a) Circuit diagram for evaluating the Thévenin impedance seen by the appended impedance Z_k. (b) Circuit diagram used to compute the current, I, conducted by Z_k. (c) The application of the substitution theorem with respect to Z_k.

The corresponding Thévenin impedance, Z_{th}, derives from a study of the test configuration of Figure 5.9, in which the independent source current, I_S, is nulled, the impedance, Z_k, in Figure 5.8 is replaced by a test current of value I_{test}, and the ratio of the resultant port voltage, V_{test}, to I_{test} is understood to be the desired Thévenin impedance. For this configuration, the network nodal admittance matrix, Y, is unchanged, but the independent network current vector, J, becomes $\mathrm{A}_{mp} I_{\text{test}}$. Thus, the resultant n-vector, E″, of nodal voltages is

$$\mathbf{E}'' = \mathbf{Y}^{-1}\mathbf{A}_{mp} I_{\text{test}} \tag{5.27}$$

and by (5.8), the voltage, V_{test}, is

$$V_{\text{test}} = \left(\mathbf{A}_{mp}^{\mathrm{T}}\mathbf{Y}^{-1}\mathbf{A}_{mp}\right) I_{\text{test}} \tag{5.28}$$

It follows that the requisite Thévenin impedance, Z_{th}, is

$$Z_{\text{th}} = \frac{V_{\text{test}}}{I_{\text{test}}} = \left(\mathbf{A}_{mp}^{\mathrm{T}}\mathbf{Y}^{-1}\mathbf{A}_{mp}\right) \tag{5.29}$$

Insofar as the appended impedance, Z_k, is concerned, the network in Figure 5.8 behaves in accordance with the circuit abstraction of Figure 5.9. The current, I, conducted by Z_k is, without approximation,

$$I = -\frac{V_{\text{th}}}{Z_{\text{th}} + Z_k} = -\left(\frac{Z_{kS}}{Z_{\text{th}} + Z_k}\right) I_S \tag{5.30}$$

where (5.26) has been used, and Z_{th} is understood to be given by (5.29). However, by the substitution theorem, the impedance, Z_k in Figure 5.8 can be supplanted by an independent current source of value I, as suggested in Figure 5.9(c). Specifically, this substitution of the impedance of interest with a current source with a value that is dictated by (5.30) guarantees that the n-vector of node voltages for the modified circuit in Figure 5.9(c) is identical to the n-vector, E', of node voltages for the topology given in Figure 5.8.

The circuit of Figure 5.9(c) now has two independent excitations: the original current sources, I_S, and the current, I, substituted for the appended impedance, Z_k. Accordingly, the current source vector for the subject circuit superimposes two current components and is given by

$$\mathbf{J} = \mathbf{A}_{12} I_S + \mathbf{A}_{mp} I = \mathbf{A}_{12} I_S - \mathbf{A}_{mp}\left(\frac{Z_{kS}}{Z_{\text{th}} + Z_k}\right) I_S \tag{5.31}$$

The corresponding vector of node voltages is, by (5.9),

$$\mathbf{E}' = \mathbf{Y}^{-1}\left[\mathbf{A}_{12} - \mathbf{A}_{mp}\left(\frac{Z_{kS}}{Z_{\text{th}} + Z_k}\right)\right] I_S \tag{5.32}$$

If analytical attention focuses on the general output voltage, $\hat{V}_{ij}$, developed between nodes i and j in the circuit of Figure 5.8,

$$\hat{V}_{ij} = \mathbf{A}_{ij}^T \mathbf{E}' \tag{5.33}$$

where

$$\hat{V}_{ij} = \left[\left(\mathbf{A}_{ij}^{\mathrm{T}} \mathbf{Y}^{-1} \mathbf{A}_{12}\right) - \left(\mathbf{A}_{ij}^{\mathrm{T}} \mathbf{Y}^{-1} \mathbf{A}_{mp}\right)\left(\frac{Z_{kS}}{Z_{\text{th}}} + Z_k\right)\right] I_S \tag{5.34}$$

The result in (5.34) is one of many possible versions of *Kron's formula*. It states that when an impedance, Z_k, is appended between nodes m and p of a linear network whose nodal admittance matrix is $\mathbf{Y}$, the perturbed voltage established between any two nodes, i and j, can be determined as a function of the parameters indigenous to the original network (prior to the inclusion of Z_k). In particular, the evaluation is executed on the original network (with Z_k absent) and exploits the original nodal admittance matrix, $\mathbf{Y}$, the original transimpedance, Z_{kS}, between the input port and the port to which Z_k is ultimately appended, and the Thévenin impedance, Z_{th}, is observed when looking into the terminal pair to which Z_k is connected.

Engineering Application of Kron's Formula

The engineering utility of Kron's formula, (5.34), is best demonstrated by examining the voltage transfer function of the network in Figure 5.8 in terms of the companion gain for the network depicted in Figure 5.5(a). Using (5.7) and noting that the perturbed output voltage is developed from node 3 to node 4, the perturbed voltage gain, $\hat{\mathrm{A}}_v$, is

$$\hat{\mathrm{A}}_v \triangleq \frac{\hat{V}_L}{V_S} = \frac{\mathbf{A}_{34}^{\mathrm{T}} \mathbf{Y}^{-1} \mathbf{A}_{12}}{Z_S} - \left(\frac{\mathbf{A}_{34}^{\mathrm{T}} \mathbf{Y}^{-1} \mathbf{A}_{mp}}{Z_S}\right)\left(\frac{Z_{kS}}{Z_{\text{th}} + Z_k}\right) \tag{5.35}$$

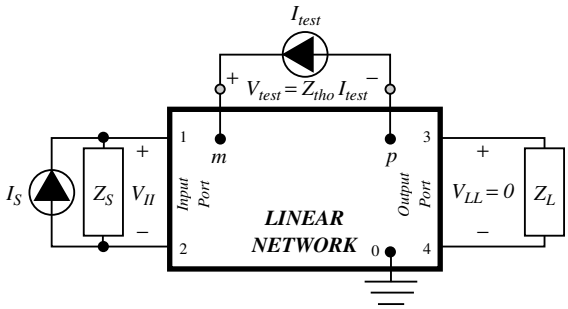


FIGURE 5.10 Network diagram pertinent to the computation of the null Thévenin impedance seen by the impedance appended to the node pair, *m-p*.

The first matrix product on the right-hand side of this relationship represents the transimpedance, Z_{LS}, from the input port to the output port of the original network, as given by (5.20). Moreover, the resultant impedance ratio, Z_{LS}/Z_S, is the voltage gain, A_v, of the original ($Z_k = \infty$) network, as delineated in (5.25). The second matrix product symbolizes the transimpedance, Z_{LK}, from the port to which the appended impedance, Z_k, is connected to the output port; that is,

$$Z_{Lk} = \mathbf{A}_{34}^{\mathrm{T}} \mathbf{Y}^{-1} \mathbf{A}_{mp} \tag{5.36}$$

Assuming $A_v \neq 0$, (5.35) can then be reduced to

$$\hat{A}_v = A_v \left[1 - \left(\frac{Z_{Lk}}{Z_{LS}} \right) \left(\frac{Z_{kS}}{Z_{\mathrm{th}} + Z_k} \right) \right] \tag{5.37}$$

This result expresses the perturbed voltage gain as a function of the original voltage gain, A_v, the input to output transimpedance, Z_{LS}, the transimpedance, Z_{kS}, from the input port to the port at which Z_k is appended, and Z_{LK}, the transimpedance from the port to which Z_k is incident to the output port. Observe that when the appended impedance is infinitely large, the perturbed gain reduces to the original voltage gain, as expected.

In an actual analytical situation, however, all of the transimpedances indicated in (5.37) need not be calculated. In order to demonstrate this contention, rewrite (5.37) in the form

$$\hat{\mathrm{A}}_v = A_v \left[\frac{1 + Y_k \left(Z_{\mathrm{th}} - \left(Z_{Lk} Z_{kS} / Z_{LS} \right) \right)}{1 + Y_k Z_{\mathrm{th}}} \right] \tag{5.38}$$

where

$$Y_k = \frac{1}{Z_k} \tag{5.39}$$

is the admittance of the appended impedance, Z_k. Now consider the test structure of Figure 5.10(a), which is the modified circuit shown in Figure 5.8, but with the appended branch supplanted by a test current source, I_{test}. With two sources, I_S and I_{test}, activating the network, superposition yields a resultant output port voltage, V_{LL}, of

$$V_{LL} = Z_{LS} I_S + Z_{LK} I_{\mathrm{test}} \tag{5.40}$$

and a test port voltage, V_{test}, of

$$V_{\text{test}} = Z_{kS} I_S - Z_{kk} I_{\text{test}} \tag{5.41}$$

For $I_S = 0$, the network in Figure 5.10(a) reduces to the configuration in Figure 5.9(a), and (5.41) delivers $V_{\text{test}}/I_{\text{test}} = Z_{kk}$. It follows that the impedance parameter, Z_{kk}, is the Thévenin impedance seen by Z_k, as determined in conjunction with an analytical consideration of Figure 5.5(a); that is, (5.41) is equivalent to the expression

$$V_{\text{test}} = Z_{kS} I_S + Z_{\text{th}} I_{\text{test}} \tag{5.42}$$

Consider the case, suggested in Figure 5.10, in which the output port voltage, V_{LL}, is constrained to zero for any and all values of the load impedance, Z_L. From (5.40), this case requires a source excitation that satisfies

$$I_S = -\left(\frac{Z_{Lk}}{Z_{LS}}\right) I_{\text{test}} \tag{5.43}$$

If this result is substituted into (5.42), the ratio, $V_{\text{test}}/I_{\text{test}}$, is found to be

$$\frac{V_{\text{test}}}{I_{\text{test}}} = Z_{th} - \frac{Z_{Lk} Z_{kS}}{Z_{LS}} \tag{5.44}$$

which mirrors the parenthesized numerator term on the right-hand side of (5.38). The ratio in (5.44) might rightfully be termed a *null Thévenin impedance*, Z_{tho}, seen by Z_k, in the sense that it is indeed the Thévenin impedance witnessed by Z_k, but under the special circumstance of a nonzero source excitation selected to *null* the output response variable of the network undergoing investigation. Thus, in Figure 5.10,

$$\left.\frac{V_{\text{test}}}{I_{\text{test}}}\right|_{\substack{\text{Source} \neq 0 \\ \text{Output} = 0}} \triangleq Z_{\text{tho}} = Z_{\text{th}} - \frac{Z_{LK} Z_{kS}}{Z_{LS}} \tag{5.45}$$

Equation (5.38) now reduces to the simpler result

$$\hat{A}_v = A_v \left(\frac{1 - Y_k Z_{\text{tho}}}{1 + Y_k Z_{\text{th}}}\right) = A_v \left(\frac{1 + \left(Z_{\text{tho}}/Z_k\right)}{1 + \left(Z_{\text{th}}/Z_k\right)}\right) \tag{5.46}$$

Equation (5.46) is both computationally useful and philosophically important. From a computational viewpoint, it allows for an efficient evaluation of the voltage transfer function of a linear network, perturbed by the addition of an impedance element between two extant nodes of the network, in terms of the voltage gain of the original, unperturbed circuit. As expected, this original voltage gain, A_v, is the voltage gain of the perturbed network for the special case of a perturbing impedance where the admittance is zero (or whose impedance value is infinitely large). Only two other parameters are required to complete the evaluation of the perturbed gain. The first is the Thévenin impedance, Z_{th}, seen by the appended impedance element. This Thévenin impedance is calculated traditionally by nulling all independent sources applied to the subject network. The second parameter is the null Thévenin impedance, Z_{tho}, which is the value of Z_{th} for the special circumstance of a test current source and independent source excitations selected to constrain the output response variable to zero. Once Z_{tho} and Z_{th} are determined, the degree to which the voltage transfer function is dependent on the appended impedance is easily determined. For example, the per-unit change in gain owing to the addition of Z_k between nodes m and p in the network of Figure 5.8 is

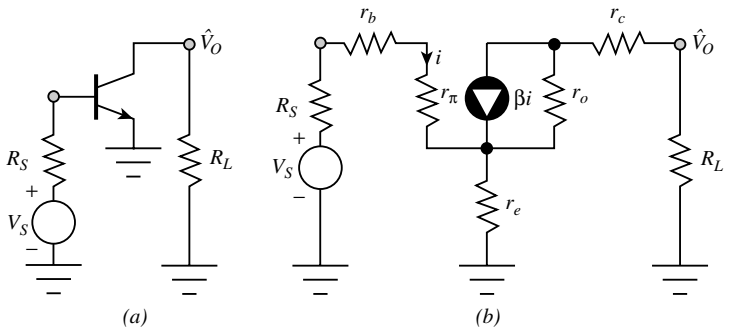


FIGURE 5.11 (a) Simplified schematic diagram of a common emitter amplifier. (b) The small signal equivalent circuit of the common emitter amplifier.

$$\frac{\Delta A_v}{A_v} = \frac{\hat{A}_v - A_v}{A_v} = \frac{Z_{\text{tho}} - Z_{\text{th}}}{Z_k + Z_{\text{th}}} \tag{5.47}$$

it is important to note that the transfer function sensitivity implied by (5.47) imposes no *a priori* restrictions on the value of Z_k. In particular, Z_k can assume any value from that of a short circuit to that of the opposite extreme of an open circuit.

From a philosophical perspective, when analytical attention focuses explicitly on feedback circuits, (5.46) can be derived from signal flow theory [10, 11] and is actually Bode's classical gain equation [12]. In the context of Bode's theory Y_k is referred to as a *reference parameter*, or a *critical parameter*, of the feedback circuit. The product, $Y_k Z_{\text{th}}$, is termed the *return ratio with respect to the critical parameter*, while the product $Y_k Z_{\text{tho}}$, is identified as the *null return ratio with respect to the cricital parameter*. Finally, when (5.46) is applied as Bode's equation, the transfer function, A_v is termed the *null transfer function*, in the sense that it is the actual transfer function of the network at hand, under the special case of a critical parameter constrained to zero.

Example 5.1. In an attempt to demonstrate the engineering utility of the foregoing theoretical disclosures, consider the problem of determining the voltage gain of the common emitter amplifier — a schematic diagram is offered in Figure 5.11(a). Without detracting from the primary intent of this example, the schematic diagram at hand has been simplified in that requisite biasing subcircuitry is not shown. Assuming that the bipolar junction transistor embedded in the amplifier operates in its linear regime, the pertinent small signal equivalent circuit is the topology depicted in Figure 5.11(b).

Assume that the amplifier source resistance, R_S, is 600 Ω and that the load resistance, R_K, is 10 kΩ. Assume further that the model parameters for the transistor are as follows: r_b (internal emitter resistance) = 2.5 Ω, r_0 (forward early resistance) = 18 kΩ, β (forward short circuit current gain) = 90, and r_c (internal collector resistance) = 70 Ω. Determine the voltage gain of the amplifier and the effect exerted on the gain by neglecting the Early resistance, R_0.

Solution.

1. Analytical simplicity traditionally dictates the tacit neglect of the forward Early resistance, r_0. This commonly invoked approximation reduces the model given in Figure 5.11(b) to the equivalent circuit in Figure 5.12(a). By inspection of the latter diagram, the approximate gain of the common emitter voltage is

$$A_v = \frac{V_O}{V_S} = -\frac{\beta R_L}{R_S + r_b + r_\pi + (\beta + 1) r_e} = -388.3 \ \text{v/v}$$

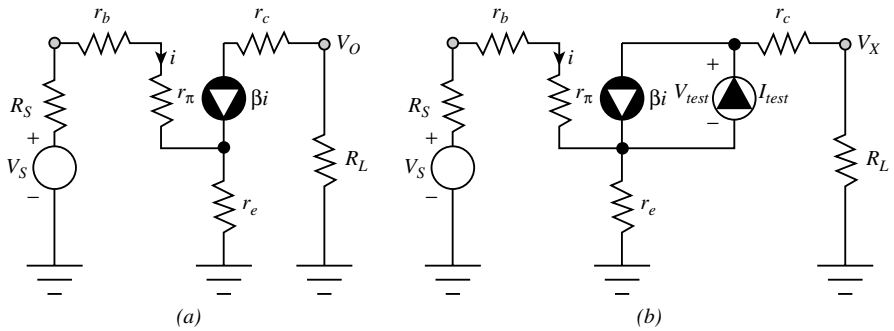


FIGURE 5.12 (a) The approximate small signal equivalent circuit of the common emitter amplifier in Figure 5.11(a). The approximation entails the tacit neglect of the forward Early resistance, r_0. (b) The test equivalent circuit used to compute the Thévenin and the null Thévenin resistances seen by r_0 in (a).

2. In order to determine the impact that r_0 has on this voltage gain, r_0 is removed from the equivalent circuit and replaced by a test current source, I_{test}, as depicted in Figure 5.12(b). With the independent input voltage, V_S, set to zero, the Thévenin resistance seen by r_0, which is the ratio, $V_{\text{test}}/I_{\text{test}}$, is easily shown to be

$$R_{\text{th}} = r_e \left\| \left(\frac{R_S + r_b + r_n}{\beta + 1} \right) + \frac{r_c + R_L}{1 + \dfrac{\beta r_e}{R_S + r_b + r_\pi + r_e}} = 9.10 \text{ k}\Omega \right.$$

On the other hand, if V_X is constrained to zero, no current flows through the load resistance branch, and therefore, I_{test} is necessarily βi. This condition gives a null Thévenin resistance of

$$R_{\text{tho}} = -\frac{r_e}{\beta} = -27.78 \times 10^{-3} \Omega$$

3. With $A_v = -388.3$ v/v, $Z_k = r_0 = 18$ kΩ, $Z_{\text{th}} = R_{\text{th}} = 9.10$ kΩ, and $Z_{\text{tho}} = R_{\text{tho}} = -27.78 \times 10^{-3'}$ Ω, (5.46) produces a corrected voltage gain of

$$\hat{A}_v = A_v \left(\frac{1 + \dfrac{R_{\text{tho}}}{r_0}}{1 + \dfrac{R_{\text{th}}}{r_0}} \right) = -258.0 \text{ v/v}$$

From (5.47), the presence of r_0, as opposed to its absence, decreases the voltage gain of the subject amplifier by almost 34%.

Example 5.2. As a second example, consider the series-shunt feedback amplifier whose schematic diagram, neglecting requisite biasing circuitry, appears in Figure 5.13(a). The analysis of this circuit is simplified by the removal of the connection of the feedback resistance, R_F, at the emitter of transistor $Q1$, as shown in Figure 5.13(b). If the voltage gain of the simplified topology is denoted as A_v, the voltage gain of the *closed loop* configuration in Figure 5.13(a) derives from (5.46), provided that the impedance, Z_k, between the indicated node pair, *m*-*p*, is taken as a short circuit; that is, $Z_k = 0$.

Assume that the amplifier source resistance, R_S, is 300 Ω, the load resistance, R_L, is 3.5 kΩ, the feedback resistance, R_F, is 1.5 kΩ, and the emitter degeneration resistance, R_{EE}, is 100 Ω. The transistor model invoked for small signal analysis is identical to that used in the preceding example, save for the proviso

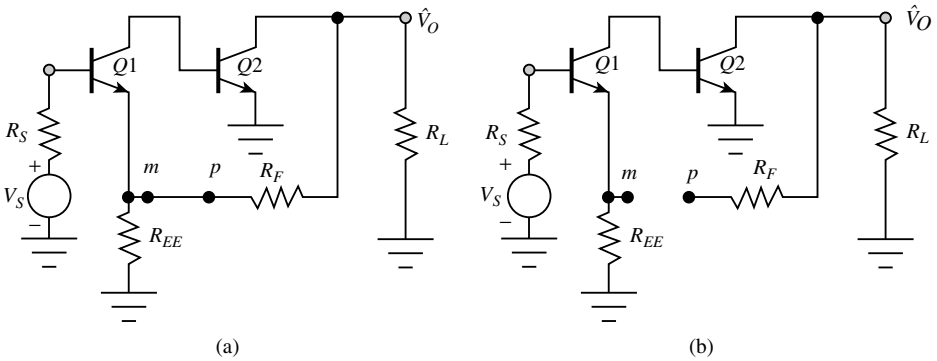


FIGURE 5.13 (a) Simplified schematic diagram of a series-shunt feedback bipolar junction transistor amplifier. The biasing subcircuitry is not shown. (b) The amplifier in (a), but with the feedback resistance connection to the emitter of transistor $Q1$ removed.

that the Early resistance, r_0, is ignored herewith. Both transistors are presumed to have identical small signal model parameters as follows: $r_b = 90\ \Omega$, $r_\pi = 1.4\ \text{k}\Omega$, $r_e = 2.5\ \Omega$, $\beta = 90$, and $r_c = 70\ \Omega$. Use Kron's theorem to determine the voltage gain of the closed loop series-shunt feedback amplifier.

Solution.

1. The voltage gain of the pertinent equivalent circuit in Figure 5.14 is straightforwardly derived as

$$A_v = \frac{V_O}{V_S} = \frac{\beta^2 R_L}{R_S + r_b + r_\pi + (\beta+1)(r_e + R_{EE})} = 2550\ \text{v/v}$$

   Observe that the feedback resistance, R_F, does not enter into this calculation because of its disconnection at the emitter of transistor $Q1$. Furthermore, the internal collector resistance, r_c, is inconsequential for both transistor stages because the neglect of the forward Early resistance, r_0, places r_c in series with a controlled current source.
2. The Thévenin resistance, R_{th}, seen by the ultimately appended short circuit between nodes m and p is now calculated through use of the model in Figure 5.14(b). For this calculation, the signal voltage, V_S, is reduced to zero. With $V_S = 0$,

$$\frac{i_1}{i_{\text{test}}} = -\frac{R_{EE}}{R_S + r_b + r_\pi + (\beta+1)(r_e + R_{EE})}$$

   Noting that $i_2 = -\beta i_1$, KVL yields

$$V_{\text{test}} = (R_{EE} + R_L + R_F) I_{\text{test}} + \left[(\beta+1)R_{EE} - \beta^2 R_L\right] i_1$$

   Using the preceding result, introducing the resistance variable, R_X, such that

$$R_X \triangleq r_e + \frac{R_S + r_b + r_\pi}{\beta+1} = 22.17\Omega$$

   and letting

$$\alpha \triangleq \frac{\beta}{\beta+1} = 0.989$$

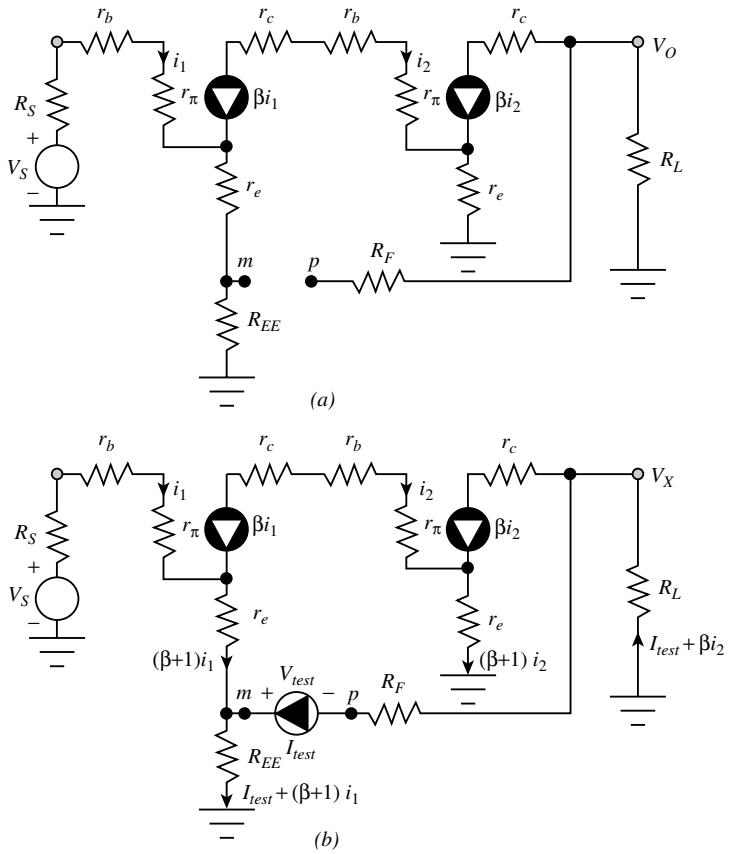


FIGURE 5.14 (a) Small signal equivalent circuit of the feedback amplifier in Figure 5.13(b). This circuit is used to compute the voltage gain with the feedback resistance disconnected at the emitter of transistor *Q*1. (b) The small signal model used to compute the Thévenin and the null Thévenin resistances seen by the short circuit that is ultimately appended to the node pair, *m*-*p*, in Figure 5.13(b).

R_{th}, which is the ratio V_{test}/I_{test}, is found to be

$$R_{th} = R_F + \left(R_{EE} \| R_X\right)\left(1 + \frac{\alpha\beta R_{EE}}{R_{EE} + R_X}\right)R_L = 260.0 \;\; K\Omega$$

3. For the evaluation of null Thévenin resistance, R_{tho}, the output voltage variable, V_X, in Figure 5.14(b) is nulled, thereby forcing the current relationship, $I_{test} = -\beta i_2 = +\beta^2 i_1$. Accordingly,

$$R_{tho} = R_F + \left(1 + \frac{1}{\alpha\beta}\right)R_{EE} = 1601\Omega$$

4. With $Z_k = 0$, $Z_{th} = R_{th} = 260$ kΩ, and $Z_{tho} = R_{tho} = 1601$ Ω, (5.46) provides, after reconnection of the feedback element, an amplifier gain of

$$\hat{A}_v = A_v\left(\frac{R_{tho}}{R_{th}}\right) = 15.70 \;\; \text{V/V}$$

It is interesting to note that, if the transistors utilized in the feedback amplifier have very large β, which is tantamount to very small R_X and $a \approx 1$, the voltage gain with the feedback element disconnected is

$$a_v \approx \frac{\beta R_L}{R_{EE}}$$

Moreover,

$$R_{\text{th}} \approx \beta R_L$$

and

$$R_{\text{tho}} \approx R_F + R_{EE}$$

It follows from (5.46) that the approximate voltage gain, subsequent to the reconnection of the feedback resistance, R_F, to the emitter of transistor $Q1$ is

$$\hat{A}_v = A_v\left(\frac{R_{\text{tho}}}{R_{\text{th}}}\right) \approx \left(\frac{\beta R_L}{R_{EE}}\right)\left(\frac{R_F + R_{EE}}{\beta R_L}\right) = 1 + \frac{R_F}{R_{EE}} = 16 \text{ V/V}$$

which is within 2% of the accurately estimated voltage gain.

Generalization of Kron's Formula

The Kron–Bode equation in (5.46) was derived expressly for investigating the voltage transfer function of a linear network to which an impedance element is appended between two network nodes. This equation also can be adapted to the problem of determining the explicit dependence of any type of transfer relationship on any parameter within any linear network.

To this end, consider any linear network, such as the generalization shown in Figure 5.15, whose, load impedance is Z_L and whose source impedance is Z_S. Identify a *critical network parameter*, say P, to which the dependence on, and sensitivity to, the overall transfer performance of the network undergoing study is of particular interest. This parameter can be, for example, a circuit branch impedance or an active element gain factor where numerical values cannot be determined accurately or controlled adequately in view of potentially unacceptable manufacturing tolerances or device fabrication uncertainties. Let the transfer function of interest be

$$H(P, Z_S, Z_L) = \frac{X_R(s)}{X_S(s)} \tag{5.48}$$

where $X_R(s)$ denotes the transform of the voltage or current response variable, and $X_S(s)$ is the transform of the voltage or current input variable. The functional notation, $H(P, Z_S, Z_L)$, underscores the observation

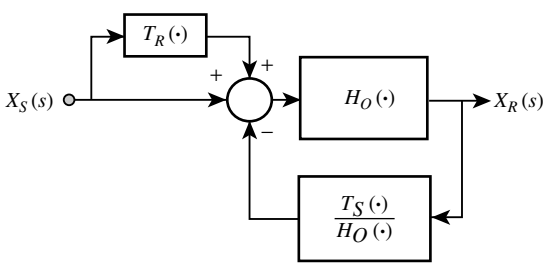


FIGURE 5.15 Generalized block diagram nodal of the I-O transfer characteristics of a linear network.

that the transfer function of the liner network is likely to be dependent on the critical parameter, P, the source impedance, Z_S, and the load impedance, Z_L. The corresponding extension of the Kron–Bode relationship is

$$H\left(P, Z_S, Z_L\right) = \frac{X_R(s)}{X_S(s)} = H\left(0, Z_S, Z_L\right)\left[\frac{1 + PQ_R\left(Z_S, Z_L\right)}{1 + PQ_S\left(Z_S, Z_L\right)}\right] \tag{5.49}$$

where $H(0, Z_S, Z_L)$, termed the *null gain or zero parameter gain,* signifies the value of the network transfer function, $H(P, Z_S, Z_L)$, when P is set to zero. This null gain must be finite and nonzero. With reference to the appended impedance formulation in (5.46), observe that the critical parameter, P, is Y_k, the admittance of the appended impedance element, while $H(0, Z_S, Z_L)$ is the gain, A_v, of the network, under the condition of an absent impedance element ($Y_k = 0$).

The product, $PQ_S\ (Z_S, Z_L)$, is termed the *return ratio with respect to parameter P,* $T_s\ (P, Z_S, Z_L)$, and the product, $PQ_R\ (Z_S, Z_L)$, is referred to as the *null return ratio with respect to P,* $T_R\ (P, Z_S, Z_L)$; that is,

$$T_R\left(P, Z_S, Z_L\right) \triangleq PQ_R\left(Z_S, Z_L\right) \tag{5.50a}$$

$$T_S\left(P, Z_S, Z_L\right) \triangleq PQ_S\left(Z_S, Z_L\right) \tag{5.50b}$$

It is to be understood that both $Q_S\ (Z_S, Z_L)$ and $Q_R\ (Z_S, Z_L)$ are independent of the critical parameter, P. With reference once again to (5.46), note that $Q_S\ (Z_S, Z_L)$ is the Thévenin impedance seen by the appended admittance, Y_k, while $Q_S\ (Z_S, Z_L)$ is the null Thévenin impedance facing Y_K.

Equation (5.49) can now be rewritten as

$$H\left(P, Z_S, Z_L\right) = \frac{X_R(s)}{X_S(s)} = H\left(0, Z_S, Z_L\right)\left[\frac{1 + T_R\left(P, Z_S, Z_L\right)}{1 + T_S\left(P, Z_S, Z_L\right)}\right] \tag{5.51}$$

Alternatively,

$$H\left(P, Z_S, Z_L\right) = \frac{X_R(s)}{X_S(s)} = H\left(0, Z_S, Z_L\right)\left[\frac{F_R\left(P, Z_S, Z_L\right)}{F_S\left(P, Z_S, Z_L\right)}\right] \tag{5.52}$$

where

$$F_S\left(P, Z_S, Z_L\right) \triangleq 1 + T_S\left(P, Z_S, Z_L\right) \tag{5.53a}$$

$$F_R\left(P, Z_S Z_L\right) \triangleq 1 + T_R\left(P, Z_S, Z_L\right) \tag{5.53b}$$

respectively, denote the *return difference with respect to P* and the *null return difference with respect to P.*

An initial appreciation of the engineering significance of the zero parameter gain, $H\ (0, Z_S, Z_L) \equiv H_0\ (\cdot)$, the return ratio, $T_S\ (P, Z_S, Z_L) \equiv T_S\ (\cdot)$, and the null return ratio, $T_R\ (P, Z_S, Z_L) \equiv T_R\ (\cdot)$, is gleaned by using (5.51) to write

$$X_R(s) = H_0(\cdot)\left[1 + T_R(\cdot)\right]X_S(s) - T_S(\cdot)X_R(s) \tag{5.54}$$

In view of the generality of the Kron–Bode formula, this algebraic manipulation of (5.51) implies that the dynamical input/output transfer relationship of all linear networks can be symbolically represented by the block diagram offered in Figure 5.15. This block diagram makes clear that because $T_S\ (\cdot)$ and $T_R\ (\cdot)$ are zero for $P = 0$, $H_0\ (\cdot)$ is the gain afforded by the network as a result of input–output electrical paths

that exclude the parameter, *P*. The diagram also suggests that $T_S(\cdot)$ is a measure of the amount of feedback incurred by parameter *P* around that part of the circuit that excludes parameter *P*. Finally, the diagram at hand implies that $T_R(\cdot)$ is a measure of the amount of *feedforward* incurred by parameter *P*. In particular, if feedback is removed, two paths remain for the transmission of a signal from the input port to the output port of a linear network. One of these paths, the transmittance of which is measured by $H_0(\cdot)$, is direct and entails the processing of the input signal by that part of the circuit that excludes *P*. The other nonfeedback path has an effective transmittance of $T_R(\cdot)\,H_0(\cdot)$. The latter path is the feedforward path in the sense that signal is processed through a signal path that is divorced from feedback and is not a result of direct source coupling through the topological part of the network that excludes parameter *P*.

Return Ratio Calculations

In the generalized transfer relationship of (5.49), the critical parameter, *P*, can assume one of only six possible forms: circuit branch admittance, circuit branch impedance, transimpedance, transadmittance, gain associated with a current-controlled current source (CCCS), and gain associated with a voltage-controlled voltage source (VCVS) [13]. The methodology underlying the computation of the return ratio and the null return ratio with respect to each of these critical parameter possibilities is given below. The case of $P = Y_k$, a circuit branch admittance, was investigated in the context of Kron's partitioning theorem. Nevertheless, it is reinvestigated next for the purpose of establishing an analytical common denominator for return ratio calculations with respect to the five other reference parameter possibilities.

Circuit Branch Admittance

Consider the network abstraction of Figure 5.16(a), which identifies a branch admittance, Y_k, as a critical parameter for analysis; that is, $P = Y_k$ in (5.49). The input excitation can be either a voltage source, or a current source and is therefore indicated as a general transformed input variable, $X_S(s)$. Similarly, the

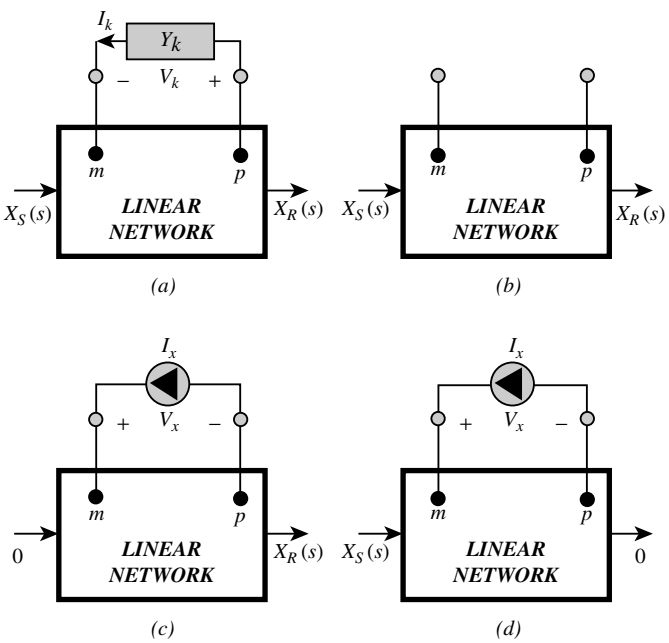


FIGURE 5.16 (a) Linear circuit for which the identified critical parameter is a branch admittance, Y_k. (b) The ratio, $X_R(s)/X_S(s)$, is the zero parameter gain $H(0, Z_S, Z_L)$. (c) The ratio, V_x/I_x, is the function $Q_S(Z_S, Z_L)$, in (5.49). (d) The ratio, V_x/I_x, is the function, $Q_R(Z_S, Z_L)$, in (5.49).

output or response variable can be either a voltage or a current, thereby encouraging the generalized transformed response notation, $X_R(s)$. The source and load impedances (or admittances) are absorbed into the network. In order to evaluate the zero parameter gain, $H(0, Z_S, Z_L)$, Y_k is set to zero by removing it from the network. An analysis is then conducted to determine the ratio $X_R(s)/X_S(s)$, of output to input variables, as suggested in Figure (5.16(b)).

As demonstrated for the case of $P = Y_k$, a circuit branch admittance, the function, $Q_S(Z_S, Z_L)$, in (5.49) is the Thévenin impedance, Z_{th}, facing Y_k. This impedance is computed by determining the ratio, V_x/I_x, with the signal source, $X_S(s)$, nulled, as indicated in Figure 5.16(c). Note in Figure 5.16(a), that the volt–ampere relationship of the branch housing Y_k is $I_k = Y_k V_k$, where the direction of the branch current, I_k, coincides with the direction of the test current source, I_x, used in the determination of Z_{th}. A comparison Figures 5.16(c) and 5.16(a) alludes to the methodology of replacing the admittance branch by a source of excitation (a current source) where the electrical nature is identical to the *dependent* electrical variable (a current, I_k) of the branch volt–ampere characteristic. Note that the polarity of the voltage, V_x, used in the determination of the test ratio, V_x/I_x, is opposite to that of the original branch voltage V_k. This is to say that although I_k and V_k are in associated reference polarity in Figure 5.16(a), I_x and V_x in the test cell of Figure 5.16(c) are in disassociated polarity.

The computation of the function, $Q_R(Z_S, Z_L)$, in (5.49) mirrors the computation of $Q_S(Z_S, Z_L)$, except for the fact that instead of setting the source excitation to zero, the output response, $X_R(s)$, is nulled. The source excitation, $X_S(s)$, remains at some computationally unimportant nonzero value, such that its effects, when superimposed over those of the test current, I_x, forces $X_R(s)$ to zero. The situation at hand is diagrammed in Figure 5.16(d).

Example 5.3. Return to the series-shunt feedback amplifier of Figure 5.29(a). Evaluate the voltage gain of the circuit, but, take the conductance, G_F, of the feedback resistance, R_F, as the critical parameter. The circuit and device model parameters remain the same as in Example 5.2: $R_S = 300\ \Omega$, $R_L = 3.5\ \text{k}\Omega$, $R_F = 1.5\ \text{k}\Omega$, $R_{EE} = 100\ \Omega$, $r_b = 90\ \Omega$, $r_\pi = 1.4\ \text{k}\Omega$, $r_e = 2.5\ \Omega$, $\beta = 90$, and $r_c = 70\ \Omega$.

Solution.

1. The zero parameter voltage gain, A_{vo}, of the subject amplifier is the voltage gain of the circuit with $G_F = 0$. But $G_F = 0$ amounts to a removal of the feedback resistance, R_F. Such removal is electrically equivalent to open circuiting the indicated node pair, *m-p*, as diagrammed in the small signal model of Figure 5.14(a). Thus, A_{vo} is identical to the gain, computed in Step (1) of Example 5.2. In particular,

$$A_{vo} = \frac{\beta^2 R_L}{R_S + r_b + r_\pi + (\beta+1)(r_e + R_{EE})} = 2550\,\text{V/V}$$

2. The model pertinent to computing the functions, $Q_S(Z_S, Z_L)$, and $Q_R(Z_S, Z_L)$, is offered in Figure 5.17. Note that the test current source, I_x, which replaces the critical conductance element, G_F, and the resultant test response voltage, V_x, are in disassociated reference polarity. As in Example 5.2, let

$$R_X \triangleq r_e + \frac{R_S + r_b + r_\pi}{\beta + 1} = 22.17$$

and

$$\alpha \triangleq \frac{\beta}{\beta + 1} = 0.989$$

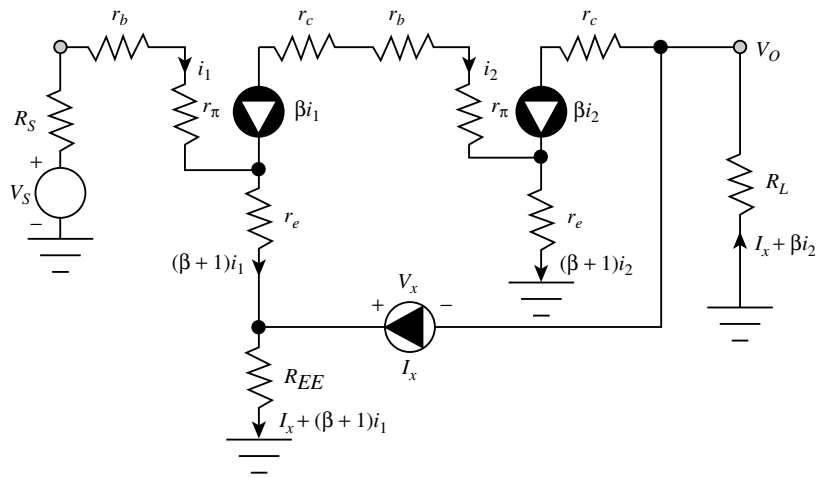


FIGURE 5.17 Circuit used to compute the return ratio and the null return ratio with respect to the conductance, G_F, in the series-shunt feedback amplifier of Figure 5.13(a).

Then, with $V_S = 0$, and writing $Q_S\ (Z_S, Z_L)$ as $Q_S\ (R_S, R_L)$ because of the lack of energy storage elements in the circuit undergoing study,

$$Q_S(R_S, R_L) = \left.\frac{V_x}{I_x}\right|_{V_S=0} = R_{EE} \| R_X + \left(1 + \frac{\alpha\beta R_{EE}}{R_{EE} + R_X}\right) R_L' = 258.5\ \text{k}\Omega$$

On the other hand,

$$Q_R(R_S, R_L) = \left.\frac{V_x}{I_x}\right|_{V_O=0} = \left(1 + \frac{1}{\alpha\beta}\right) R_{EE} = 101.1\ \Omega$$

3. Substituting the preceding results into (5.49), and recalling that $G_F = 1/R_F$, the voltage gain of the series-shunt feedback amplifier is found to be

$$A_v = \frac{V_O}{V_S} = A_{vo}\left[\frac{1 + \dfrac{Q_R(R_S, R_L)}{R_F}}{1 + \dfrac{Q_S(R_S, R_L)}{R_F}}\right] = 15.7\ \text{V/V}$$

which is the gain result deduced previously.

Circuit Branch Impedance

In the circuit of Figure 5.18(a), a branch impedance, Z_k, is selected as a critical parameter for analysis; that is $P = Z_k$ in (5.49). The zero parameter gain, $H(0, Z_S, Z_L)$, is evaluated by replacing Z_k with a short circuit, as suggested in Figure 5.18(b).

The volt–ampere characteristic equation of the critical impedance branch is $V_k = Z_k I_k$, where, of course, the branch voltage, V_k, and the branch current, I_k, are in associated reference polarity. Because the dependent variable in this volt–ampere expression is a branch voltage, the return and null return ratios are calculated by replacing the subject branch impedance by a test voltage source, V_x. As suggested in Figure 5.18(c), the ratio, I_x/V_x, under the condition of nulled independent sources, gives the function $Q_S\ (Z_S, Z_L)$ in (5.49). On the other hand, and as depicted in Figure 5.18(d), the ratio I_x/V_x, with a nulled

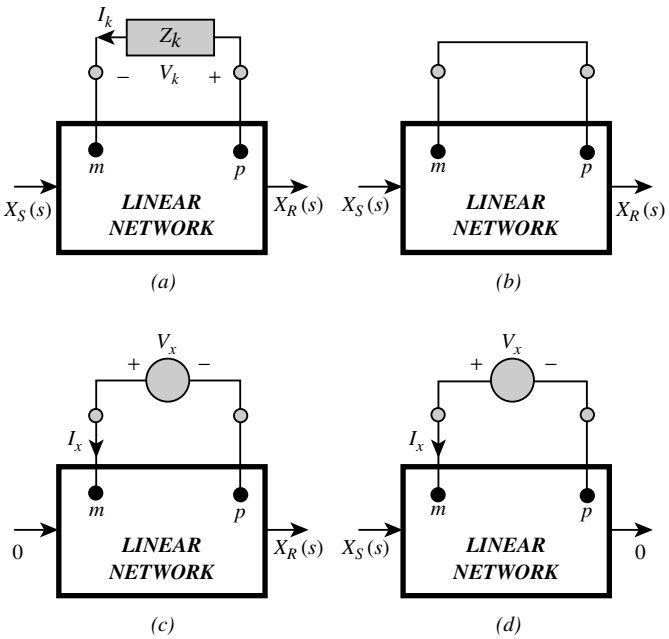


FIGURE 5.18 (a) Linear circuit for which the identified critical parameter is a branch impedance, Z_k. (b) The ratio, $X_R(s)/X_S(s)$, is the zero parameter gain, $H(0, Z_S, Z_L)$. (c) The ratio, I_x/V_x, is the function, $Q_S(Z_S, Z_L)$, in (5.49). (d) The ratio, I_x/V_x, is the function, $Q_R(Z_S, Z_L)$, in (5.49).

response, yields $Q_R(Z_S, Z_L)$. Observe that in the present situation, the functions, $Q_S(Z_S, Z_L)$ and $Q_R(Z_S, Z_L)$ are, respectively, the Thévenin and the null Thévenin admittances facing the branch impedance, Z_k.

Circuit Transimpedance

In the circuit of Figure 5.19(a), a circuit transimpedance, Z_t, is selected as a critical parameter for analysis; that is, $P = Z_t$ in (5.49). The zero parameter gain, $H(0, Z_S, Z_L)$, is evaluated by replacing the current-controlled voltage source (CCVS) by a short circuit, as shown in Figure 5.19(b).

The volt–ampere characteristic equation of the critical transimpedance branch is $V_k = Z_t I_k$, where I_k is the controlling current for the controlled source branch. Because the dependent variable in this volt–ampere expression is a branch voltage, the return and null return ratios are calculated by replacing the CCVS with a test voltage source, V_x. However, as indicated in Figures 5.19(c) and (d), the polarity of V_x mirrors that of the voltage, V_k, developed across the controlled branch. With I_x taken as a current flowing in the controlling branch in a direction opposite to the polarity of the original controlling current, I_k, the ratio, I_x/V_x, under the condition of nulled independent sources, gives the function, $Q_S(Z_S, Z_L)$ in (5.49). On the other hand, and as depicted in Figure 5.19(d), the ratio, I_x/V_x, with a nulled response, yields $Q_R(Z_S, Z_L)$.

Circuit Transadmittance

In the network of Figure 5.20(a), a circuit transadmittance, Y_t, is selected as the critical parameter. The zero parameter gain, $H(0, Z_S, Z_L)$, is evaluated by replacing the voltage-controlled current source (VCCS) with an open circuit, as shown in Figure 5.20(b).

The volt–ampere characteristic question of the critical transadmittance branch is $I_k = Y_t V_k$, where V_k is the controlling voltage for the VCCS. Because the dependent variable in this volt-ampere expression is a branch current, the return and null return ratios are calculated by replacing the VCCS with a test current source, I_x, where, as indicated in Figures 5.20(c) and (d), the polarity of I_x mirrors that of the current, I_k, flowing through the controlled branch. With V_x taken as a voltage developed across the

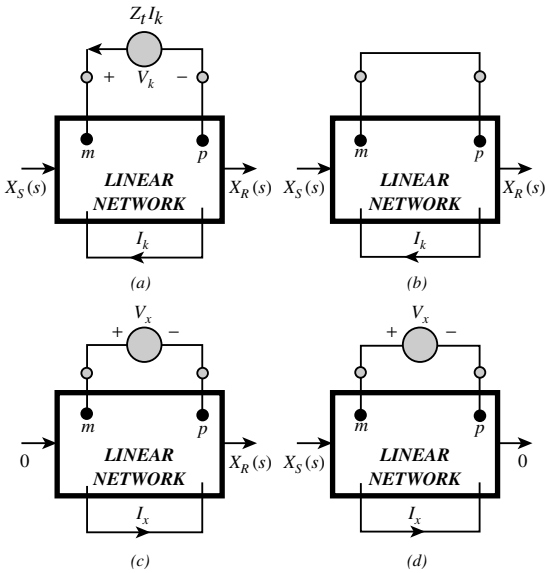


FIGURE 5.19 (a) Linear circuit for which the identified critical parameter is a circuit transimpedance, Z_t. (b) The ratio, $X_R(s)/X_S(s)$, is the zero parameter gain, $H(0, Z_S, Z_L)$. (c) The ratio, I_x/V_x, is the function, $Q_S(Z_S, Z_L)$, in (5.49). (d) The ratio, I_x/V_x, is the function, $Q_R(Z_S, Z_L)$, in (5.49).

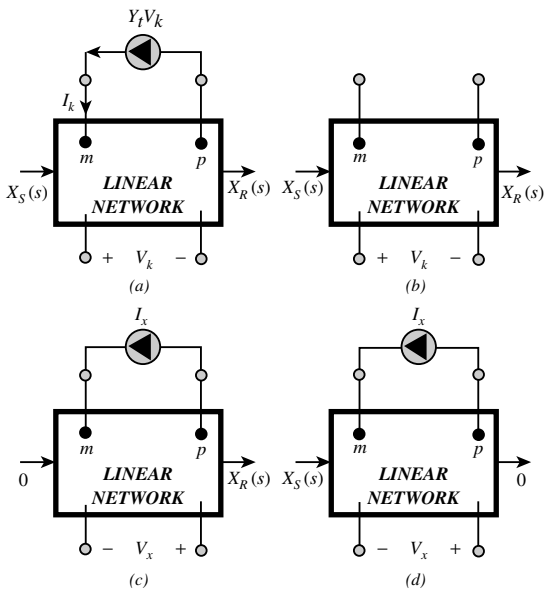


FIGURE 5.20 (a) Linear circuit for which the identified critical parameter is a circuit transadmittance, Y_t, (b) The ratio, $X_R(s)/X_S(s)$, is the zero parameter gain, $H(0, Z_S, Z_L)$. (c) The ratio, V_x/I_x, is the function $Q_S(Z_S, Z_L)$, in (5.49). (d) The ratio, V_x/I_x, is the function $Q_R(Z_S, Z_L)$, in (5.49).

controlling branch in a direction opposite to the polarity of the original controlling voltage, V_k, the ratio, V_x/I_x, under the condition of nulled independent sources, gives the function, $Q_S(Z_S, Z_L)$ in (5.43). On the other hand, and as offered in Figure 5.20(d), the ratio, V_x/I_x, under the condition of a nulled response, yields $Q_R(Z_S, Z_L)$.

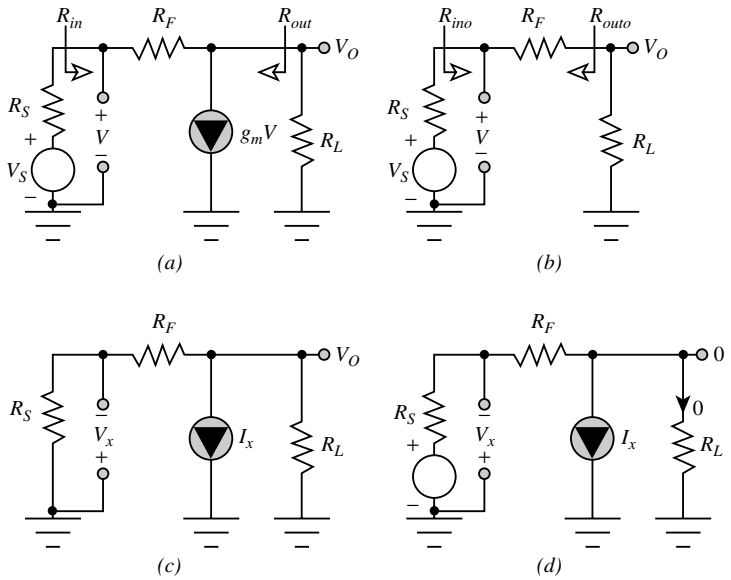


FIGURE 5.21 (a) The low-frequency, small-signal model of a voltage feedback amplifier. (b) The circuit used to evaluate the zero parameter ($g_m = 0$) gain. (c) The circuit used to evaluate the return ratio with respect to g_m. (d) The circuit used to computer the null return ratio with respect to g_m.

Example 5.4. The circuit in Figure 5.21(a) is a low-frequency, small-signal model of a voltage feedback amplifier. With the transconductance, g_m, selected as the reference parameter of interest, derive a general expression for the voltage gain, $A_v = V_O/V_S$. Approximate the final result for the special case of very large g_m.

Solution.

1. The zero parameter voltage gain, A_{vo}, derives from an analysis of the circuit structure given in Figure 5.21(b). The diagram differs from Figure 5.21(a) in that the current conducted by the controlled source branch has been nulled by open circuiting said branch. By inspection of the subject model,

$$A_{vo} = \frac{R_L}{R_L + R_F + R_S}$$

2. The diagram given in Figure 5.21(c) is appropriate to the computation of the return ratio, $T_S(g_m, Z_S, Z_L) = g_m Q_S(R_S, R_L)$ with respect to the critical transconductance, g_m. A comparison of the model at hand with the diagram in Figure 5.21(a) confirms that the controlled source, $g_m V$, is replaced by an independent current source, I_x, which flows in a direction identical to that of the controlled source it supplants. The ratio, V_x/I_x, is to be computed, where V_x is developed, antiphase to V, across the branch that supports the original controlling voltage for the VCCS. A straightforward analysis produces

$$Q_S\left(R_S, R_L\right) = \frac{V_x}{I_x}\left(\frac{R_L}{R_L + R_F + R_S}\right)R_S \equiv A_{vo}R_S$$

3. The null return ratio, $T_R(g_m, Z_S, Z_L) = g_m Q_R(R_S, R_L)$, with respect to g_m is obtained from an analysis of the circuit in Figure 5.21(d). Observe a nulled output voltage, with zero current flow through the load resistance, R_L. Observe further that the signal source voltage is nonzero. The specific value of this source voltage is not crucial, and is therefore not delineated. An analysis reveals

$$Q_R(R_S, R_L) = \frac{V_x}{I_x} = -R_F$$

4. Using (5.49), the voltage gain of the circuit undergoing study is found to be

$$A_v = \frac{V_O}{V_S} = A_{vo}\left[\frac{1 + g_m(-R_F)}{1 + g_m R_S A_{vo}}\right]$$

which, for large g_m, reduces to

$$A_v \approx -\frac{R_F}{R_S}$$

Gain of a Current-Controlled Current Source

For the network in Figure 5.22(a), the reference parameter is α_k, the gain associated with a CCCS. The zero parameter gain, $H(0, Z_S, Z_L)$, is evaluated by replacing this CCCS with an open circuit, as depicted in Figure 5.22(b).

The volt–ampere characteristic equation of the branch in which the reference parameter is embedded is $I_k = \alpha_k I_j$, where I_j is the controlling current for the CCCS. Because the dependent variable in this volt-ampere characteristic is a branch current, the return and null return ratios are calculated by replacing the CCCS with a test current source, I_x. As indicated in Figures 5.22(c) and (d), the polarity for I_x mirrors that of the current, I_k, flowing through the controlled branch. Let I_y be the resultant current conducted by the controlling branch, and let this current flow a direction opposite to the polarity of the original controlling current. Then, the current ratio, I_y/I_x, computed under the condition of nulled independent sources, is the function, $Q_S(Z_S, Z_L)$ in (5.49). Similarly, and as suggested in Figure 5.22(d), the ratio, I_y/I_x, under the condition of a nulled response, yields $Q_R(Z_S, Z_L)$.

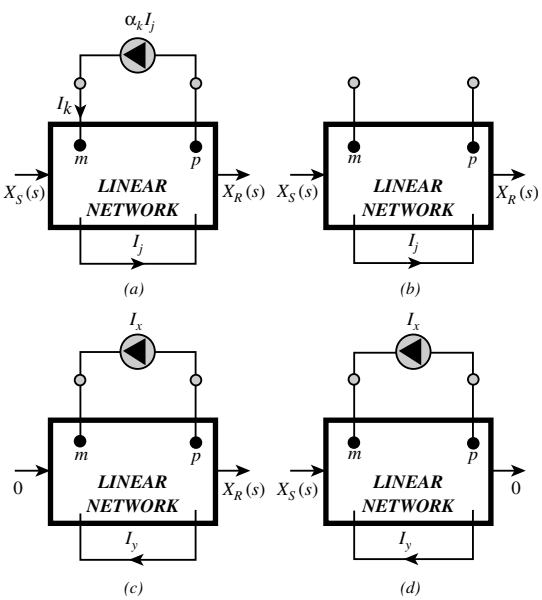


FIGURE 5.22 (a) Linear circuit for which the identified critical parameter is the current gain α_k associated with a CCCS. (b) The ratio, $X_R(s)/X_S(s)$, is the zero parameter gain, $H(0, Z_S, Z_L)$. (c) The ratio, I_y/I_x, is the function, $Q_S(Z_S, Z_L)$, in (5.49). (d) The ratio, I_y/I_x, is the function, $Q_R(Z_S, Z_L)$, in (5.49).

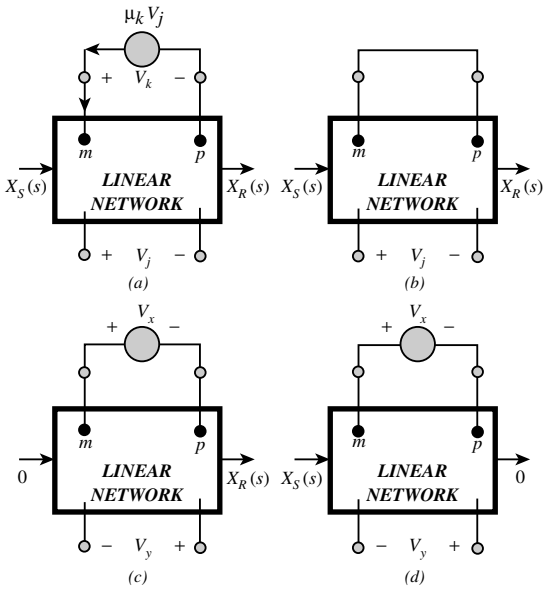


FIGURE 5.23 (a) Linear circuit for which the identified critical parameter is the voltage gain, μ_{k1} associated with a VCVS. (b) The ratio, $X_R(s)/X_S(s)$, is the zero parameter gain, $H(0, Z_S, Z_L)$. (c) The ratio, V_y/V_x, is the function $Q_S(Z_S, Z_L)$, in (5.49). (d) The ratio, V_y/V_x, is the function, $Q_R(Z_S, Z_L)$, in (5.49).

Gain of a Voltage-Controlled Voltage Source

In the network of Figure 5.23(a), the selected reference parameter is μ_k, the gain corresponding to a VCVS. The zero parameter gain, $H(0, Z_S, Z_L)$, is evaluated by replacing this VCVS with a short circuit, as per Figure 5.23(b).

The volt–ampere characteristic equation of the dependent generator branch is $V_k = \mu_k V_j$, where V_j is the controlling voltage for the VCVS. Because the dependent variable in this volt–ampere expression is a branch voltage, the return and null return ratios are calculated by replacing the VCVS with a test voltage source, V_x, where as indicated in Figures 5.23(c) and (d), the polarity of V_x is identical to that of the voltage, V_k, developed across the controlled branch. Let V_y be the resultant voltage established across the controlling branch, and let the polarity of this voltage be in a direction opposite to that of the original controlling voltage. Then, the voltage ratio, V_y/V_x, computed under the condition of nulled independent sources, is the function, $Q_S(Z_S, Z_L)$, in (5.49). As suggested in Figure 5.23(d), the voltage ratio, V_y/V_x, under the condition of a nulled response, yields $Q_R(Z_S, Z_L)$.

Evaluation of Driving Point Impedances

Having formulated generalized techniques for computing the return ratio and the null return ratio with respect to any of the six possible types of critical circuit parameters, the application of (5.49) is established as a powerful and computationally expedient vehicle for evaluating any transfer function of any linear network. The only restriction limiting the utility of (5.49) is that parameter P must be selected in such a way as to ensure that the zero parameter transfer function is finite and nonzero.

Equation (5.49) is commonly used to evaluate the voltage gain, current gain, transimpedance gain, or transadmittance gain of feedback and other types of complex circuitry. However, the expression is equally well suited to determining the driving point input impedance seen by the source impedance, as well as the driving point output impedance seen by the terminating load impedance. In fact, once the return ratios relevant to the gain of interest are found, these I-O impedances can be determined with minimal additional analysis.

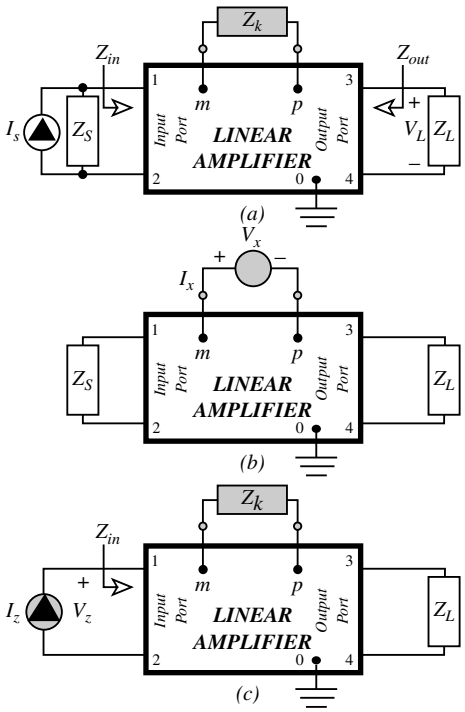


FIGURE 5.24 (a) A liner amplifier for which the input impedance is to be determined. (b) The circuit used for calculating the return ratio with respect to Z_k. (c) The circuit used for calculating the driving point input impedance.

Without loss of generality, the foregoing contention is explicitly demonstrated in conjunction with a transimpedance amplifier whose reference parameter is selected to be a branch impedance, Z_k. To this end, consider the circuit abstracted in Figure 5.24, for which the driving point input impedance, Z_{in}, is to be determined. The input excitation is a current, I_S, and in response to this input, a signal voltage, V_L, is developed across the load impedance, Z_L. Using (5.49), the I-O transimpedance, $Z_T(Z_k, Z_S, Z_L)$, is

$$Z_T(Z_K, Z_S, Z_L) = \frac{V_L(s)}{I_S(s)} = Z_T(0, Z_S, Z_L)\left[\frac{1 + Z_k Q_R(Z_S, Z_L)}{1 + Z_k Q_S(Z_s, Z_L)}\right] \tag{5.55}$$

where $Z_T(0, Z_S, Z_L)$ is the circuit transimpedance for $Z_k = 0$, $Z_k\, Q_R\,(Z_S, Z_L)$ is the null return ratio with respect to Z_k, and $Z_K\, Q_S\,(Z_S, Z_L)$ is the return ratio with respect to Z_k. For future reference, the circuit appropriate to the calculation of the function, $Q_S\,(Z_S, Z_L)$, is drawn in Figure 5.24(b).

The input impedance derives from an analytical consideration of the cell depicted in Figure 5.24(c), in which the Norton representation of the signal source has been supplanted by a test current source of value I_z. Note that the load impedance remains as the terminating element for the output port. The transfer relationship of interest is the ratio, V_z/I_z, which is the desired driving point input impedance, Z_{in}. Taking care to choose Z_k, the reference parameter for the gain enumeration, as the reference parameter for the input impedance determination, (5.49) gives

$$Z_{\text{in}} = \frac{V_z}{I_z} = Z_{\text{ino}}\left[\frac{1 + Z_k Q_{RR}(Z_S, Z_L)}{1 + Z_k Q_{SS}(Z_S, Z_L)}\right] \tag{5.56}$$

In this expression, Z_{ino} generally derives straightforwardly because it is the $Z_k = 0$ value of Z_{in}; that is, Z_{in} is evaluated for the special case of a nulled reference parameter. Such a null in the present situation is equivalent to short-circuiting Z_k, as indicated in Figure 5.25(a). The function, $Q_{SS}\,(Z_S, Z_L)$, is the delineated

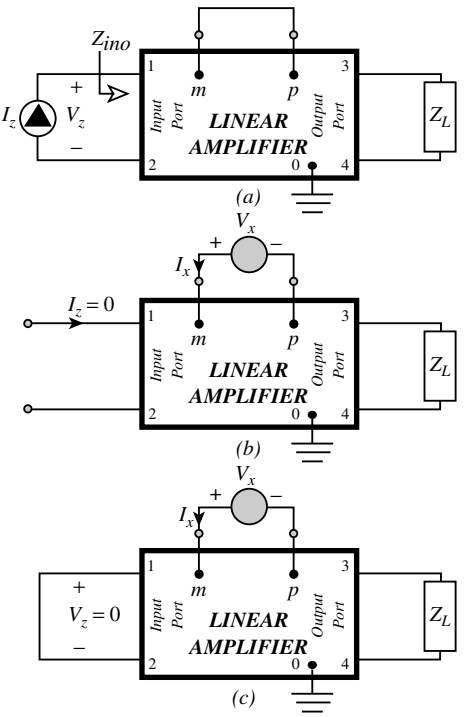


FIGURE 5.25 (a) The circuit used to evaluate the zero parameter driving point input impedance. (b) The computation, relative to input impedance, of the return ratio with respect to Z_k. (c) The computation, relative to input impedance, of the null return ratio with respect to Z_k.

I_x/V_x ratio, for the case of a source excitation (I_z in the present case) set to zero. The pertinent circuit diagram is the structure in Figure 5.25(b). This last circuit differs from the circuit, shown in Figure 5.24(b), exploited to find $Q_S(Z_S, Z_L)$ in the gain relationship of (5.50) in only one way: Z_S has been removed, and thus, effectively, Z_S has been set to an infinitely large value. It follows that

$$Q_{SS}(Z_S, Z_L) \equiv Q_S(\infty, Z_L) \tag{5.57}$$

In other words, a circuit analysis aimed toward determining $Q_{SS}(Z_S, Z_L)$ is unnecessary. Instead, $Q_{SS}(Z_S, Z_L)$ is found by evaluating $Q_S(Z_S, Z_L)$, which is already known from the gain analysis, *at* $Z_S = \infty$.

To evaluate $Q_{RR}(Z_S, Z_L)$, the foregoing I_x/V_x ratio is calculated for the case of zero response. In the present situation the response is the voltage, V_z, and accordingly, the appropriate circuit is depicted in Figure 5.25(c). However, a comparison of the circuit at hand with the structure in Figure 5.24, which is exploited to evaluate the return ratio in the gain equation, indicates that it differs only in that Z_S is now constrained to zero to ensure $V_z = 0$. It is therefore apparent that in (5.56)

$$Q_{RR}(Z_S, Z_L) \equiv Q_S(0, Z_L) \tag{5.58}$$

Equation (5.56) is now expressible as

$$Z_{\text{in}} = \frac{V_z}{I_z} = Z_{\text{ino}}\left[\frac{1 + Z_k Q_S(0, Z_L)}{1 + Z_k Q_S(\infty, Z_L)}\right] \tag{5.59}$$

which is occasionally referred to as Blackman's formula [14].

Analogous considerations at the output port in the circuit of Figure 5.24(a) dictate a driving point output impedance, Z_{out}, of

$$Z_{out} = Z_{outo}\left[\frac{1+Z_kQ_S(Z_S,0)}{1+Z_kQ_S(Z_S,\infty)}\right] \tag{5.60}$$

where, similar to Z_{ino}, Z_{outo}, the $Z_k = 0$ value of Z_{out}, must be finite and nonzero. Although the preceding two relationships were derived for the case in which the selected reference parameter is a branch impedance, both expressions are applicable for any reference parameter, P. In general,

$$Z_{in} = Z_{ino}\left[\frac{1+PQ_s(0,Z_L)}{1+PQ_S(\infty,Z_L)}\right] \tag{5.61a}$$

$$Z_{out} = Z_{outo}\left[\frac{1+PQ_s(Z_S,0)}{1+PQ_S(Z_S,\infty)}\right] \tag{5.61b}$$

Example 5.5. Use the pertinent results of Example 5.4 to derive expression for the driving point input resistance, R_{in}, and the driving point output resistance, R_{out}, of the feedback amplifier in Figure 5.21(a).

Solution.

1. With g_m set to zero, an inspection of the circuit diagram in Figure 5.21(b) delivers

$$R_{ino} = R_F + R_L$$

$$R_{outo} = R_F + R_S$$

2. From the second step in the solution to Example 5.4, the function, $Q_S(R_S, R_L)$, to which the return ratio, $T_S(g_m, R_S, R_L)$ is directly proportional, was found to be

$$Q_S(R_S,R_L) = \left(\frac{R_L}{R_L+R_F+R_S}\right)R_S$$

It follows that

$$Q_S(0,R_L) = 0$$

$$Q_S(\infty,R_L) = R_L$$

Moreover,

$$Q_S(R_S,0) = 0$$

$$Q_S(R_S,\infty) = R_S$$

3. Equations (5.61a) and (b) resultantly yield

$$R_{in} = R_{ino}\left[\frac{1+g_mQ_S(0,Z_L)}{1+g_mQ_S(\infty,Z_L)}\right] = \frac{R_R+R_L}{1+g_mR_L}$$

for the driving point input resistance and

$$R_{\text{out}} = R_{\text{out}\,0}\left[\frac{1+g_m Q_S\left(R_S,0\right)}{1+g_m Q_S\left(R_S,\infty\right)}\right] = \frac{R_F + R_S}{1+g_m R_S}$$

for the driving point output resistance.

Sensitivity Analysis

Yet another advantage of the Kron–Bode formula is its amenability to evaluating the impact exerted on a circuit transfer relationship by potentially large fluctuations in the reference parameter *P*. This convenience stems from the fact that parameter *P* is isolated in (5.49); that *is*, $H\,(0, Z_S, Z_L)$, $Q_R\,(Z_S, Z_L)$, and (Z_S, Z_L) are each independent of *P*. A quantification of this impact is achieved by exploiting the *sensitivity function*,

$$S_P^H \triangleq \frac{\Delta H/H}{\Delta P/P} \tag{5.62}$$

which compares the per unit change in transfer function, $\Delta H/H$, resulting from a specified per unit change $\Delta P/P$ in a critical parameter. In particular, the notation in this definition is such that if *H* designates the transfer characteristic, $H\,(P_0, Z_S, Z_L)$, at the nominal parameter setting, $P = P_0$, $(H + \Delta H)$ signifies the perturbed characteristic, $H(P_0 + \Delta P, Z_S, Z_L)$ where P_0 is altered by an amount ΔP_0. Using (5.49) and dropping the functional notation in (5.53a) and (5.53b), it can be demonstrated that

$$S_P^H \triangleq \frac{F_R - F_S}{F_R\left[F_S + \left(F_S - 1\right)\left(\dfrac{\Delta P}{P_0}\right)\right]} \tag{5.63}$$

where F_S and F_R are understood to be evaluated at the nominal parameter setting, $P = P_0$. It should be emphasized that unlike a more traditional sensitivity analysis, such as that predicated on the adjoint network [15], (5.63) is easy to use manually and does not rely on an *a priori* assumption of small parametric changes.

References

[1] G. Kron, *Tensor Analysis of Networks*, New York: John Wiley & Sons, 1939.

[2] G. Kron, "A set of principals to interconnect the solutions of physical systems," *J. Appl. Phys.*, vol. 24, pp. 965–980, Aug. 1953.

[3] F. H. Branin, Jr., "The relation between Kron's method and the classical methods of circuit analysis," *IRE Conv. Rec.*, part 2, pp. 3–28, 1959.

[4] F. H. Branin, Jr., "A sparse matrix modification of Kron's method of piecewise analysis," *Proc. IEEE Int. Symp. Circuits Systems*, pp. 383–386, 1975.

[5] R. A. Rohrer, "Circuit partitioning simplified," *IEEE Trans. Circuits Syst.*, vol. 35, pp. 2–5, Jan. 1988.

[6] N. B. Rabbat and H. Y. Hsieh, "A latent macromodular approach to large-scale sparse networks," *IEEE Trans. Circuits Syst.*, vol. CAS-23, pp. 745–752, Dec. 1976.

[7] J. Choma, Jr., "Signal flow analysis of feedback networks," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 455–463, April 1990.

[8] N. B. Rabbat, A. L. Sangiovanni-Vincentelli, and H. Y. Hsieh, "A multilevel Newton algorithm with macromodeling and latency for the analysis of large-scale nonlinear circuits in the time domain," *IEEE Trans. Circuits Syst.*, vol. CAS-26, pp. 733–741, Sep. 1979.

[9] N. Balabanian and T. A. Bickart, *Electrical Network Theory,* New York: John Wiley & Sons, 1969, pp. 73–77.
[10] S. J. Mason, "Feedback theory—some properties of signal flow graphs," *Proc. IRE*, vol. 41, pp. 1144–1156, Sep. 1953.
[11] S. J. Mason, "Feedback theory—further properties of signal flow graphs," *Proc. IRE*, vol. 44, pp. 920–962, July 1956.
[12] H. W. Bode, *Network Analysis and Feedback Amplifier Design*, New York: D. Van Nostrand Company, 1945 (reprinted, 1953), chap. 4
[13] J. Choma, Jr., *Electrical Networks: Theory and Analysis*, New York: Wiley Interscience, 1985, pp. 590–598.
[14] S. K. Mitra, *Analysis and Synthesis of Linear Active Networks*, New York: John Wiley & Sons, 1969, p. 206.
[15] S. W. Director and R. A. Rohrer, "The generalized adjoint network and network sensitivities," *IEEE Trans. Circuit Theory*, vol. CT-16, pp. 318–328, 1969.

6

Tableau and Modified Nodal Formulations

Jiri Vlach
University of Waterloo, Canada

6.1 Introduction

Network analysis is based on formulation of the relevant equations and on their solutions. Various approaches are possible. If we wish to get as much theoretical information as possible, we may resort to hand analysis and keep the elements as variables (literal parameters). In such cases, it is an absolute necessity to use a method that leads to the smallest possible number of equations. If we plan to use a computer, then we can accept methods which lead to larger systems, but the methods must be relatively easy to program. The purpose of this chapter is to give an overview of the various possibilities and point out advantages and disadvantages. Many more details are available in [1, 2].

Section 6.2 is a summary of the nodal and mesh formulations. We review them because they are the best ones for analysis of small networks. Tableau formulation, given in Section 6.3, is very general, but requires special solution routines, probably not available to the reader. Section 6.4 describes the best method for computerized solutions; it is used in many commercial simulators. If nonlinear elements are involved, then iterative solution methods must be used; an introduction on how to deal with nonlinear elements is given in Section 6.5. Finally, Section 6.6 presents a method that is suitable for hand solutions of active networks and which automatically leads to the smallest system of equations.

6.2 Nodal and Mesh Formulations

Classical methods use two types of network equations formulation: the nodal and the mesh. The first one is based on Kirchhoff's current law (KCL): the sum of currents flowing away from a node is equal to zero. The mesh method is based on Kirchhoff's voltage law (KVL): the sum of voltages around any loop is equal to zero.

For simple problems, both methods are about equivalent, but nodal formulation is more general. The mesh formulation is suitable only for planar networks: It must be possible to draw the network without any element crossing over any other element. Many practical networks are planar, but it is not always easy to see that it is actually the case.

We first introduce some definitions. A positive current flows from a terminal with a higher potential to a terminal with a lower potential. This is sketched on the two-terminal element in Figure 6.1. If we

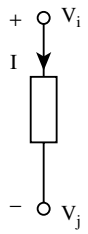


FIGURE 6.1 Definition of positive current direction with respect to the voltage across the element.

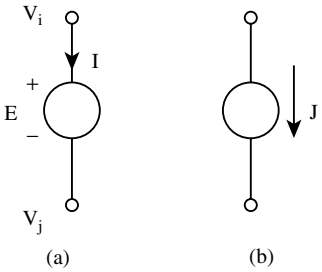


FIGURE 6.2 (a) Direction of positive current through an independent voltage source. (b) Direction of the current through an independent current source.

use this definition, then the product of the current and of the voltage across the element, $V_i - V_j$, expresses the power consumed by the element. If the current flows in opposite direction, then the element is delivering power.

We will use the previous definition of a positive current for *all* elements of the network, irrespective of what their role eventually may be. Thus, a positive current through an independent voltage source will flow from the more positive terminal to the less positive terminal, as sketched in Figure 6.2(a). The current flowing through an independent current source is indicated on its symbol, Figure 6.2(b), but the voltage across it is not defined; it depends on the network.

The nodal formulation uses the principle that the sum of currents at any node must be equal to zero at any instant of time. To apply this rule in an efficient way, we realize that before we solve the equations, we do not know which way the currents will actually flow. All we know is that *if* a node is more positive than all the other nodes, then all currents must flow *away* from this node.

In nodal formulation, the unknowns are nodal voltages and the equations express the sum of currents flowing away from the node. To write the equations we use element admittances: in Laplace transform $Y_C = sC$ for a capacitor, $Y_L = 1/sL$ for an inductor, and $Y_G = G = 1/R$ for a resistor. It is advantageous to use G, because we thus avoid fractions in the equations.

We demonstrate how to set up the equations by considering the network in Figure 6.3. The nodal voltages are denoted V_1 and V_2 and ground (the lower line) is considered to be at zero potential. We do not know which of these nodes is more positive, but we can *assume* that any node we consider at a given moment is the most positive one. This assumption has the consequence that all currents must flow *away*

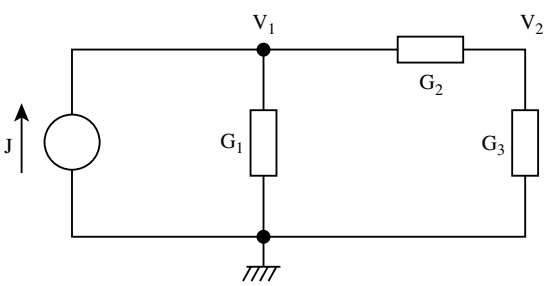


FIGURE 6.3 Example for nodal formulation.

from this node. For the given network, the current through G_1 will flow down and its value will be $I_{G1} = G_1V_1$, current through G_2 will flow from left to right and will be $I_{G2} = G_2(V_1 - V_2)$. Current from the independent current source flows *into* the node and thus must be subtracted. Together, the sum of the currents at node 1 will be zero:

$$G_1V_1 + G_2(V_1 - V_2) - J = 0$$

Moving to the second node, we still do not know which node is more positive, but we can still make the assumption that *now* it is *this* node that is the most positive one. If we make such an assumption, then all currents must flow away from this node: The current through G_3 will be $I_{G3} = G_3V_2$and will flow down, the current through G_2 will flow from right to left and will be $I_{G2} = G_2(V_2 - V_1)$. In this expression, the first voltage within the parentheses must be the assumed higher potential. This is expressed by the equation

$$G_3V_2 + G_2(V_2 - V_1) = 0$$

It is advantageous to put the equations into matrix form, with the known independent source transferred to the right side:

$$\begin{bmatrix} G_1 + G_2 & -G_2 \\ -G_2 & G_2 + G_3 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \end{bmatrix} \begin{bmatrix} J \\ 0 \end{bmatrix}$$

The preceding steps were simple because we selected elements that can be handled by this formulation. Unfortunately, many practical elements are not expressed in terms of currents. For instance, a voltage source connected between nodes *i* and *j*, with its positive reference on node *i*, is described by the equation

$$V_i - V_j = E$$

A positive current does flow through such an element from *i* to *j*, but is not available in its defining equation. In fact, all voltage sources, independent or dependent, will create this problem. Another element which cannot be handled directly is a short circuit. It is described by the equation

$$V_i - V_j = 0$$

and current is not a part of its definition.

We can always use transformations by applying various theorems such as the Thévenin and Norton transformations or source splitting, and eventually arrive at a network in which all elements have voltage as the independent variable. Such transformations are practical for hand analysis, but are not advantageous for computerized solutions. This is the reason why other formulations have been invented.

Consider next the mesh equations where we use the KVL and impedances of the elements: $Z_L = sL$ for the inductor, $Z_C = 1/sC$ for the capacitor, and R for the resistor. In this formulation, we sum the voltages across the elements in a given closed loop. Because this method is suitable for planar networks only, we usually use the concept or circulating mesh currents, indicated on the network in Figure 6.4. The currents $I_1$and I_2 create voltage drops across the resistors. When considering the first mesh, we take the current I_1 as a positive one. The voltage across R_1 is $V_{R1} = R_1I_1$. The voltage across R_2 is $V_{R2} = R_2(I_1 - I_2)$ and the voltage source contributes a value E to the equation. According to our earlier definition, a positive current flows from plus to minus, but I_1 actually goes in the opposite direction through the voltage source. Thus, the voltage across E must be taken with a negative sign and the sum of voltages around the first mesh is

$$R_1I_1 + R_2(I_1 - I_2) - E = 0$$

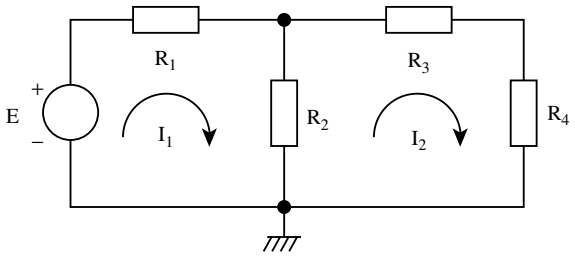


FIGURE 6.4 Example for mesh formulation.

When we move to the second mesh, we consider the current I_2 as positive and the sum of voltage drops around the second mesh is

$$R_2(I_2 - I_1) + R_3 I_2 + R_4 I_2 = 0$$

The equations can be collected into one matrix equation

$$\begin{bmatrix} R_1 + R_2 & -R_2 \\ -R_2 & R_2 + R_3 + R_4 \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \end{bmatrix} \begin{bmatrix} E \\ 0 \end{bmatrix}$$

Each of these fundamental formulations has its problems.

In nodal formulation, we can deal directly with the following elements:

Current source, J
Conductance, $G = 1/R$
Capacitor admittance, sC
Voltage controlled current source, VC
Inductor admittance, $1/sL$

In mesh formulation, we can deal directly with the elements:

Voltage source, E
Resistor, R
Inductor impedance, sL
Current controlled voltage source, CV
Capacitor impedance, $1/sC$.

All other elements create problems and must be dealt with by the Thévenin and Norton theorems and/or source splitting.

As an example, we take the network in Figure 6.5(a). It is directly suitable for mesh formulation, but we demonstrate both. For simplicity, all resistors have unit values.

The mesh formulation, with the indicated circulating currents I_1 and I_2, leads to the equations

$$(R_1 + R_2) I_1 - R_2 I_2 = E$$

$$-R_2 I_1 + (R_2 + R_3) I_2 + 3I_1 = 0$$

Inserting numerical values

$$2I_1 - I_2 = E$$

$$2I_1 + 2I_2 = 0$$

The solution is $I_1 = E/3$, $I_2 = -E/3$ and $V_1 = R_2(I_1 - I_2) = 2E/3$.

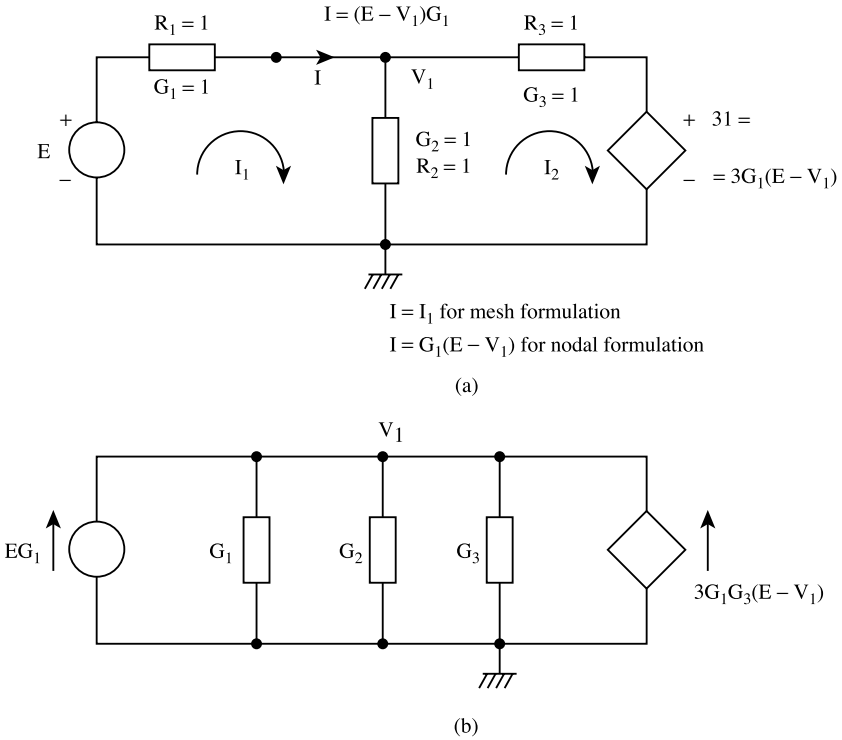


FIGURE 6.5 (a) Example of a network suitable for mesh formulation. (b) Modification of the network in Figure 6.5(a) to be suitable for nodal formulation.

To use nodal formulation, we must apply several transformation steps. First, we must express the controlling currents as $I = G_1(E - V_1) = E - V_1$ and replace I in the definition of the current controlled voltage source. This has been done in the figure. Afterward, applying Thévenin–Norton transformation we change the voltage sources, in series with resistors, into current sources, in parallel with the same resistors. We get the network in Figure 6.5(b). It has only one node for which the balance of current is

$$\left(G_1 + G_2 + G_3\right)V_1 - EG_1 - 3G_1G_3\left(E - V_1\right) = 0$$

Inserting numerical values and solving, we get the same V_1 as previously.

If we have a mixture of elements, such transformations will always be lengthy, will require redrawings, and can lead to errors. To reduce the chance of such errors to a minimum, in computer applications we need formulations that avoid transformations and use descriptions of the elements as they are given. This is done in both the tableau and nodal formulations, the subjects of the following sections.

Although the nodal and mesh formulations are not always easy to apply, we must stress that they are the best ones for hand solutions. They may require several steps of transformations and redrawings, but ultimately they lead to the smallest possible systems of equations.

6.3 Graphs and Tableau Formulation

Tableau is the most general formulation because the solution simultaneously provides the voltages across all elements, the currents through all elements, and all nodal voltages. The difficulty is that tableau leads to such large systems of equations that complicated sparse matrix solvers are an absolute necessity. Most readers will not have access to such routines, therefore, we will explain its properties only to the extent necessary for understanding. We advise the reader not to use it.

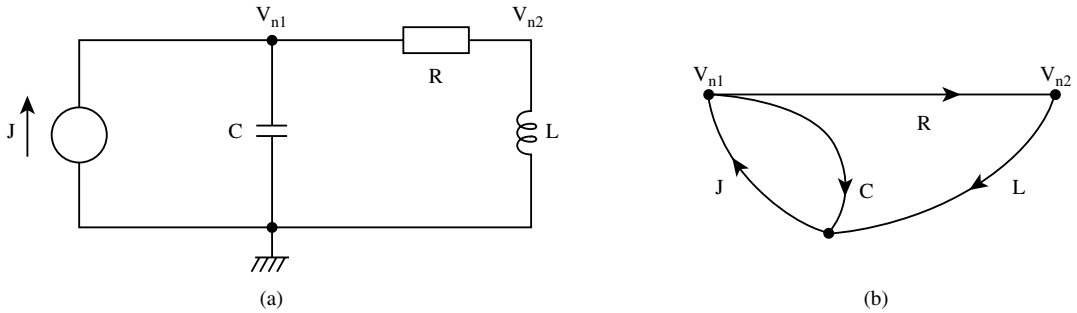


FIGURE 6.6 (a) A simple network; (b) its graph.

Tableau formulation needs for its construction the concept of graphs and the concept of the incidence matrix. Consider the network in Figure 6.6(a). A graph of the network replaces each element by a line. We will use oriented graphs, with arrows, because they can be identified with the flow of currents. In all passive elements, the current can flow in any direction and the orientation of the graph is entirely our choice. We do not have such freedom when we consider sources. The direction of the current through the current source is given by the arrow marked at its symbol and we use the same direction in the graph. For the voltage source, the direction of the graph will be from plus to minus, in agreement with our previous explanations. Each node is marked by the node voltage and the line representing the element is given the name of the element. Following these rules, we have constructed the graph in Figure 6.6(b). Because the directions of the arrows are the assumed directions of currents, we can write KCL for the two nodes: positive direction is away from the node, negative is into the node. The sums of currents for the nodes are

$$-I_J + I_C + I_R = 0$$

$$-I_R + I_L = 0$$

This can also be summarized in one matrix equation

$$\begin{bmatrix} -1 & 1 & 1 & 0 \\ 0 & 0 & -1 & 1 \end{bmatrix} \begin{bmatrix} I_j \\ I_C \\ I_R \\ I_L \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \end{bmatrix}$$

or

$$\mathbf{AI} = 0 \tag{6.1}$$

The matrix A is called the *incidence* matrix. It has as many rows as there are ungrounded nodes, and as many columns as the number of elements. Note that +1 in any given row indicates that we expect the current to flow away from the node, –1 means the opposite.

Still more information can be extracted from this matrix. Denote the nodal voltages by subscripts n in V_{n1} and V_{n2}, as done in Figure 6.6. The voltages across the elements will have as subscripts the names of the elements. We can write the following set of equations which couple the voltages across the elements with the nodal voltages:

$$V_J = -V_{n1}$$

$$V_C = V_{n1}$$

$$V_R = V_{n1} - V_{n2}$$
$$V_L = V_{n2}$$

In matrix form, this is equivalent to

$$\begin{bmatrix} V_J \\ V_C \\ V_R \\ V_L \end{bmatrix} = \begin{bmatrix} -1 & 0 \\ 1 & 0 \\ 1 & -1 \\ 0 & 1 \end{bmatrix} \begin{bmatrix} V_{n1} \\ V_{n2} \end{bmatrix}$$

The matrix is the transpose of the incidence matrix and we can generalize

$$\mathbf{V}_{el} - \mathbf{A}^I \mathbf{V}_n = 0 \tag{6.2}$$

Complete formulation needs expressions that couple the element currents and the element voltages. Writing them in the same sequence as for the graph, and using Laplace transformation, we have

$$I_J = J$$
$$I_C = sCV_C$$
$$V_R = RI_R$$
$$V_L = sLI_L$$

For matrix notation, we need an expression that, with a proper choice of entries, will cover all possible elements. Such an expression is

$$YV_{el} + ZI_{el} = W$$

For instance, if we consider the current source, we set $Y = 0$, $Z = 1$ and $W = J$, which gives the preceding equation. Similar choices can be made for the other elements.

The KCL equation, $\mathbf{AI} = \mathbf{0}$, the KVL equation, $\mathbf{V}_{el} - \mathbf{A}^T\mathbf{V}_n = \mathbf{0}$, and the previous equation $\mathbf{YV}_{el} + \mathbf{ZI}_{el} = \mathbf{W}$ are collected in one matrix equation. Any sequence can be used; we have chosen

$$\begin{aligned} \mathbf{V}_{el} - \mathbf{A}^T\mathbf{V}_n &= 0 \\ \mathbf{YV}_{el} + \mathbf{ZI}_{el} &= \mathbf{W} \\ \mathbf{AI}_{el} &= 0 \end{aligned} \tag{6.3}$$

and in matrix form

$$\begin{bmatrix} \mathbf{1} & \mathbf{0} & -\mathbf{A}^{\mathrm{T}} \\ \mathbf{Y} & \mathbf{Z} & \mathbf{0} \\ \mathbf{0} & \mathbf{A} & \mathbf{0} \end{bmatrix} \begin{bmatrix} \mathbf{V}_{\mathrm{el}} \\ \mathbf{I}_{\mathrm{el}} \\ \mathbf{V}_{\mathrm{n}} \end{bmatrix} = \begin{bmatrix} \mathbf{0} \\ \mathbf{W} \\ \mathbf{0} \end{bmatrix} \tag{6.4}$$

Once the incidence matrix is available, writing this matrix equation is actually quite simple. First determine its size: it will be twice the number of elements plus the number of nodes. For our example, it will be 10. The system equation is in Figure 6.7 where all zero entries were omitted to clearly show the

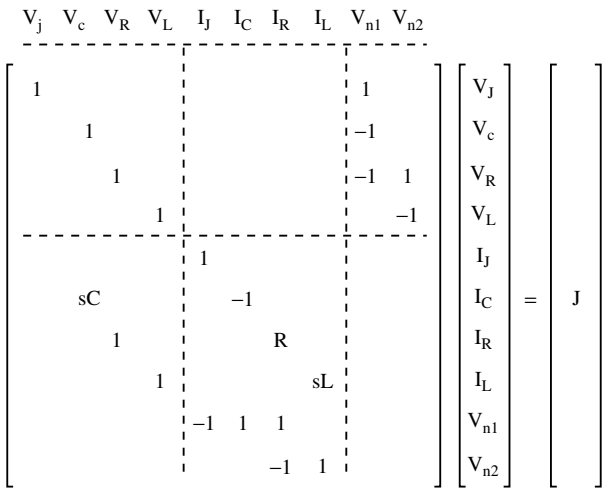


FIGURE 6.7 Tableau formulation for the network in Figure 6.6.

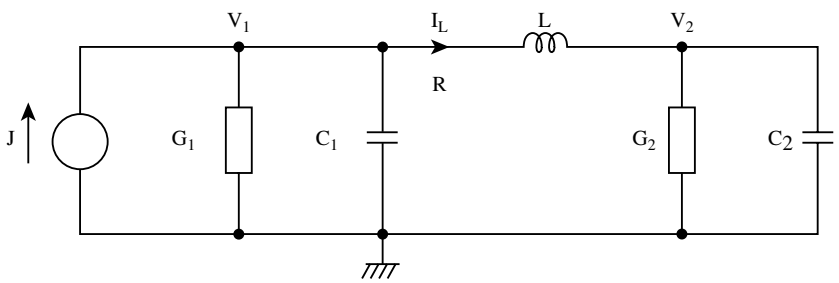


FIGURE 6.8 Example for modified nodal formulation.

structure. In the top partition is a unit matrix and the negative of the transpose of the **A** matrix. In the bottom partition is the incidence matrix, **A**. The middle portion is filled, element by element, using the previous element equation. For better understanding, it is a good idea to write the variables above the matrix, as shown, because each column of the matrix is multiplied by the variable which appears above it.

We have used this simple example to point out the main difficulty of the tableau formulation: the system becomes very large. In nodal formulation, this problem would lead to only two equations. However, the tableau system matrix has many zeros and is said to be sparse. Sparse systems are always solved by special routines which, roughly speaking, do not store the zeros and do not operate on them. Such codes are quite difficult to write, and in tableau we have the additional difficulty that the matrix has a complicated structure. We discussed this formulation more as a warning instead of a recommendation. Unless a suitable sparse matrix solver is already available, this formulation should be avoided.

6.4 Modified Nodal Formulation

Modified nodal formulation is an extension of the nodal formulation and is the method of choice for computerized analysis. It is used in most commercial simulators, and we will explain it in considerable detail.

When nodal formulation is taught in schools, inductors are usually taken as admittances, $Y_L = 1/sL$. This is fine, as long as we work by hand and derive the network function. For instance, nodal equations for the network in Figure 6.8 would be written in the form

$$\left(G_1 + sC_1 + \frac{1}{sL}\right) V_1 - \frac{1}{sL} V_2 = J$$

$$-\frac{1}{sL} V_1 + \left(G_2 + sC_2 + \frac{1}{sL}\right) V_2 = 0$$

However, this creates a problem for computerized solutions. Multiplication by s represents differentiation in the Laplace transform and $1/s$ represents integration. As a result, the two equations are actually a set of two integro-differential equations, and we do not normally have methods to solve them directly in such a form. In all computerized methods, we use integration of systems of first-order differential equations and the preceding equations cannot be arranged into such a form. What we need is a method which will keep all frequency-dependent elements in the form sC or sL, with the variable s in the numerator. Such possibility exists if we take into account a new variable, the current through the inductor. Writing KCL for the two nodes of Figure 6.8

$$\left(G_1 + sC_1\right) V_1 + I_L = J$$

$$\left(G_2 + sC_2\right) V_2 - I_L = 0$$

We now have two equations but three variables. What we have not used yet is an expression which couples the voltages across the inductor with the current through it. The relationship is $V_1 - V_2 = sLI_L$, but because we do not know any of these three variables, we transfer everything to the left and write the last equation

$$V_1 - V_2 = sLI_L = 0 \tag{6.5}$$

All three can be put into a matrix form

$$\begin{bmatrix} G_1 + sC_1 & 0 & 1 \\ 0 & G_2 + sC_2 & -1 \\ 1 & -1 & -sL \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ I_L \end{bmatrix} = \begin{bmatrix} J \\ 0 \\ 0 \end{bmatrix}$$

In this equation, the nodal portion is the 2×2 matrix in the upper left corner and information about the inductor is collected in the right-most column and the lowest row. A larger network will have a larger nodal portion, but we still increase the matrix by one row and column for each inductor. This can be prepared as a general *stamp* as shown in Figure 6.9. The previous matrix is the empty box, separated by the dashed lines, the voltages and the current above the stamp indicate the variables relevant to the inductor, while on the left the letters i and j give the rows (node numbers) where L is connected. Should we have a network with two inductors, we add one row and one column for each.

The next element we take is the voltage-controlled current source; it was already mentioned is Section 6.2 as an element which can be taken into the nodal formulation. The *VC* is shown in Figure 6.10.

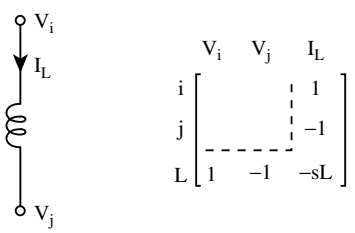


FIGURE 6.9 Stamp of an inductor.

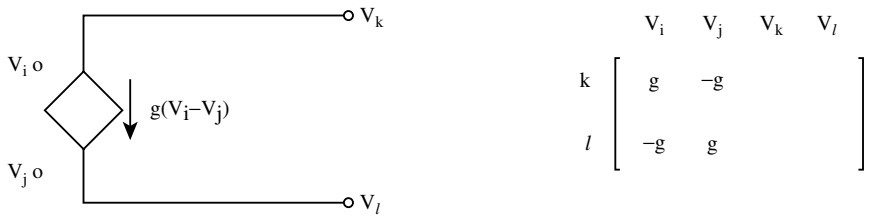


FIGURE 6.10 Stamp for a voltage-controlled current source.

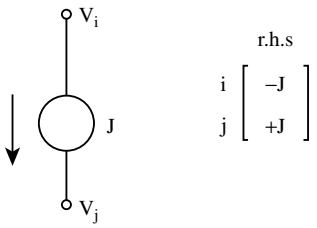


FIGURE 6.11 Stamp for an independent current source.

It adds the current $g(V_i - V_j)$ to node k and subtracts the same current at node l. It permits us to write the stamp, Figure 6.10. The element influences the balance of currents at node i and j, and the variables which multiply its transconductance are V_i and V_j.

Network theory defines two independent sources: the current and the voltage source. The current source can be taken into consideration in the right-hand side (r.h.s.) of the nodal portion; its stamp is in Figure 6.11. The independent voltage source cannot be taken directly and we must add its current as a new variable. Consider the source with a resistor in series, shown in Figure 6.12. The voltage relationships are

$$V_i - V_j - rI_E = E \tag{6.6}$$

The current I_E adds to the balance of currents at node i, because it flows away from it. It is subtracted from the balance of currents at node j. The current is taken as a new variable and the equation is attached to previous equations. The stamp is as shown. It is, in fact, a combined stamp for several elements. If we set $r = 0$, we have an ideal voltage source. If we set both r and E equal to zero, we have a stamp for a short circuit. For better understanding, consider the example in Figure 6.13. The network has two ungrounded nodes for which we can write the nodal equations:

$$\left(G_1 + sC_1\right) V_1 + I_E = 0$$

$$\left(G_2 + sC_2\right) V_2 - I_E = 0$$

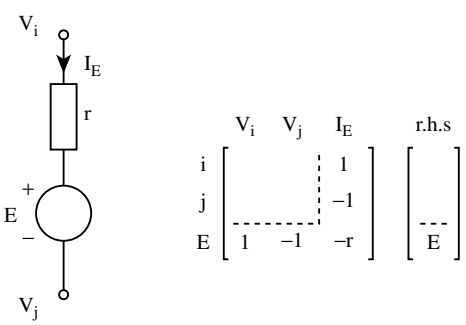


FIGURE 6.12 Stamp for an independent voltage source. The resistor value can be zero.

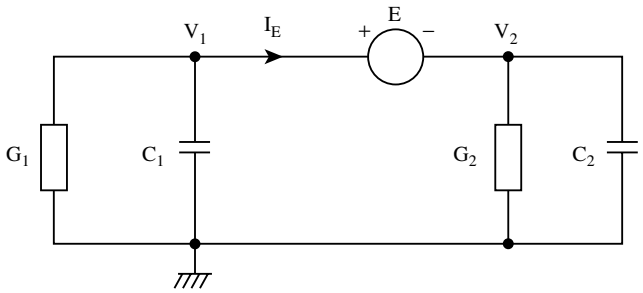


FIGURE 6.13 Example with a floating voltage source.

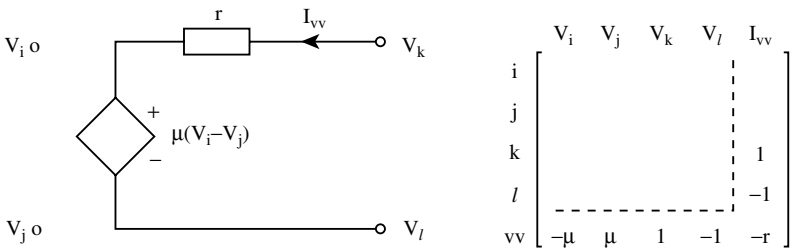


FIGURE 6.14 Stamp for a voltage-controlled voltage source. The resistor value can be zero.

To these equations, we must add the equation describing the properties of the ideal voltage source,

$$V_i - V_j = E \tag{6.7}$$

Collecting into matrix form

$$\begin{bmatrix} G_1 + sC_1 & 0 & 1 \\ 0 & G_2 + sC_2 & -1 \\ 1 & -1 & 0 \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ I_E \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ E \end{bmatrix}$$

We still need stamps for the remaining three dependent sources. The voltage controlled voltage source, *VV*, is shown in Figure 6.14. For generality, we added the internal resistor. The output is described by the equation

$$\mu\left(V_i - V_j\right) + rI_{VV} = V_k - V_l \tag{6.8}$$

None of the voltages is known, so we transfer everything to the left side. Input terminals do not influence the balance of currents, but the output terminals do. At node *k* we must add I_{VV}, and subtract the same at node *l*. The stamp is in Figure 6.14.

The current controlled current source, *CC*, is shown in Figure 6.15. The input terminals are short circuited and

$$V_i - V_j = 0 \tag{6.9}$$

No information is available about the output voltages, but we know that the current is α times the input current, and thus only one additional variable is needed. Balances of currents are influenced at all four nodes: positive *I* at node *i*, negative at node *j*, positive current αI at node *k*, and negative αI at node *l*.

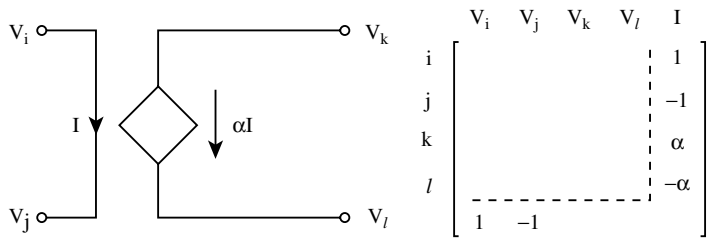


FIGURE 6.15 Stamp for a current-controlled current source.

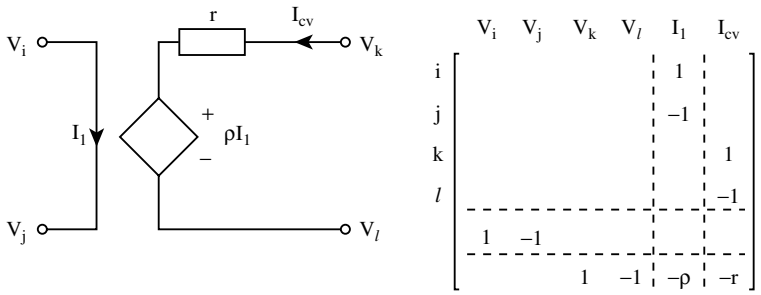


FIGURE 6.16 Stamp for a current-controlled voltage source. The resistor value can be zero.

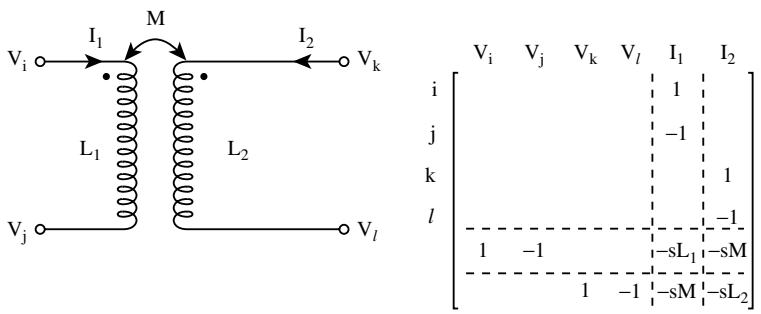


FIGURE 6.17 Stamp for a transformer.

The stamp is in Figure 6.15, where the added column takes care of the currents and the additional row describes the properties of the short circuit.

The most complicated dependent source is the current-controlled voltage source, *CV*, shown in Figure 6.16. We can consider it as a combination of a short circuit and a voltage source. The equation for the short circuit is the same as for the *CC*. The output is defined by the equation

$$V_k - V_l = \rho I_1 + rI_{CV} \tag{6.10}$$

and because none of the variables is known, we transfer everything to the left. This element adds two rows and two columns to the previously defined matrix. Its stamp is in Figure 6.16. As before, the internal resistor r can be set equal to zero.

Modified nodal formulation easily takes into account a transformer (see Figure 6.17). It is described by the equations

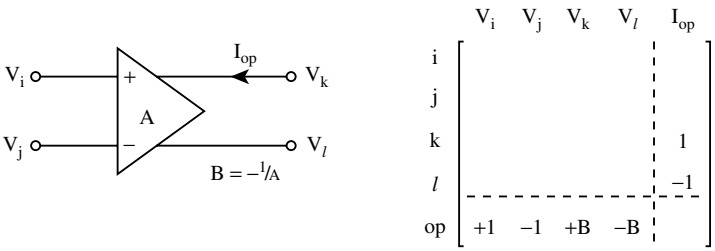


FIGURE 6.18 Stamp for an ideal operational amplifier.

$$\begin{aligned} V_i - V_j &= sL_1 I_1 + sMI_2 \\ V_k - V_l &= sL_1 I_1 + sL_2 I_2 \end{aligned} \tag{6.11}$$

None of the variables is known, and we transfer everything to the left. The currents influence the balance of currents at all nodes. The stamp of the transformer is in Figure 6.17.

The last element we consider is an ideal operational amplifier. It is sometimes taken as a voltage-controlled voltage source with very high gain, but it is preferable to have a stamp that can take into account ideal properties as well. The element is shown in Figure 6.18. The terminal l is usually grounded, but we will keep it floating to make the stamp more general. No current flows into the device at the input terminals. The output equation is

$$V_k - V_l = A\left(V_i - V_j\right) \tag{6.12}$$

Because a computer cannot handle infinity, it is advantageous to introduce the inverted gain

$$B = -1/A \tag{6.13}$$

and modify the previous equation to

$$V_i - V_j + BV_k - BV_l = 0 \tag{6.14}$$

This equation is attached to the set of equations and the balance of currents is influenced at nodes k and l. This leads to the stamp in Figure 6.18. If we set $B = 0$, the operational amplifier becomes ideal, with no approximation.

An example will show how the stamps are used. Consider the network in Figure 6.19. It has no practical application, but serves well for the demonstration of how to set up the modified nodal matrix. A short circuit, indicating the controlling current of the current-controlled voltage source is taken into account

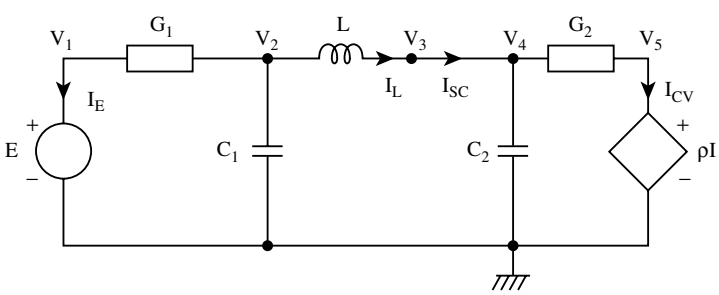


FIGURE 6.19 Example showing the use of modified nodal formulation.

by increasing the number of nodes. The network has five nongrounded nodes, and thus the dimension of its nodal portion will be 5. The voltage source will increase the system matrix by one row and one column, the inductor also, and the *CV* will need two more rows and columns; altogether the matrix will be 9×9. Write first the nodal portion by disregarding entirely the other elements. This creates the upper left partition. Using the stamps we add first the voltage source, then the inductor, next the short circuit, and finally the current-controlled voltage source. The system matrix is

$$\left[\begin{array}{ccccc|c|c|cc}
G_1 & -G_1 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \\
-G_1 & G_1+sC_1 & 0 & 0 & 0 & 0 & 1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -1 & 1 & 0 \\
0 & 0 & 0 & sC_2+G_2 & -G_2 & 0 & 0 & -1 & 0 \\
0 & 0 & 0 & -G_2 & G_2 & 0 & 0 & 0 & 1 \\
\hline
1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
\hline
0 & 1 & -1 & 0 & 0 & 0 & -sL & 0 & 0 \\
\hline
0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & -\rho & 0
\end{array}\right]
\left[\begin{array}{c} V_1 \\ V_2 \\ V_3 \\ V_4 \\ V_5 \\ I_E \\ I_L \\ I_{SC} \\ I_{CV} \end{array}\right]
=
\left[\begin{array}{c} 0 \\ 0 \\ 0 \\ 0 \\ 0 \\ E \\ 0 \\ 0 \\ 0 \end{array}\right]$$

Modified nodal formulation is the most important method for computer applications. The reader can find additional information in the books [1, 2].

6.5 Nonlinear Elements

In previous sections, we used the Laplace transform to explain the various methods of formulation. Because we dealt with linear elements, the systems of equations were linear and it was possible to cast them into matrix forms.

If we must consider nonlinear elements, we face many restrictions. The Laplace transform cannot be used. Various concepts based on it, like the network functions, the poles and the zeros, cannot be applied. Only two types of analysis are available:

The dc solution (operating point)
Time-domain solution for a given input signal

Once we have nonlinear elements, we cannot write the equations in matrix form; all we can do is write KCL equations. We must also find another method for the solution of such nonlinear equations.

Consider two differentiable equations in two unknowns

$$\begin{aligned} f_1(v_1, v_2) &= 0 \\ f_2(v_1, v_2) &= 0 \end{aligned} \tag{6.15}$$

The functions can be expanded into Taylor series and the series truncated after the linear terms:

$$f_1(v_1 + \Delta v_1,\ v_2 + \Delta v_2) = f_1(v_1,\ v_2) + \frac{\partial f_1}{\partial v_1}\Delta v_1 + \frac{\partial f_1}{\partial v_2}\Delta v_2 + \cdots = 0$$

$$f_2(v_1 + \Delta v_1,\ v_2 + \Delta v_2) = f_2(v_1,\ v_2) + \frac{\partial f_2}{\partial v_1}\Delta v_1 + \frac{\partial f_2}{\partial v_2}\Delta v_2 + \cdots = 0$$

We are not considering higher-order terms, so the equation will not be exactly zero, but we can still try to find Δv_1 and Δv_2. Transferring the known function values to the right

$$\frac{\partial f_1}{\partial v_1}\Delta v_1 + \frac{\partial f_1}{\partial v_2}\Delta v_2 = -f_1(v_1, v_2)$$
$$\frac{\partial f_2}{\partial v_1}\Delta v_1 + \frac{\partial f_2}{\partial v_2}\Delta v_2 = -f_2(v_1, v_2)$$

This is a system of *linear* equations, and we can rewrite it in matrix form

$$\begin{bmatrix} \frac{\partial f_1}{\partial v_1} & \frac{\partial f_1}{\partial v_2} \\ \frac{\partial f_2}{\partial v_1} & \frac{\partial f_2}{\partial v_2} \end{bmatrix} \begin{bmatrix} \Delta v_1 \\ \Delta v_2 \end{bmatrix} = \begin{bmatrix} -f_1(v_1, v_2) \\ -f_2(v_1, v_2) \end{bmatrix} \tag{6.16}$$

The matrix on the left is called the *Jacobian*; on the right is the negative of the functions. Once this *linear* system is solved, we can get new values of the variables by writing

$$v_1^{(i+1)} = v_1^{(i)} + \Delta v_1^{(i)}$$
$$v_2^{(i+1)} = v_2^{(i)} + \Delta v_2^{(i)} \tag{6.17}$$

In this equation, we added the superscript to indicate iteration. The process is repeated until all Δv_i become sufficiently small. This iterative method is usually referred to as the Newton–Raphson iteration and is written in the form

$$\mathbf{J}(\mathbf{v}^{(i)})\ \Delta\mathbf{v}^{(i)} = -\mathbf{f}(\mathbf{v}^{(i)})$$
$$\mathbf{v}^{(i+1)} = \mathbf{v}^{(i)} + \Delta\mathbf{v}^{(i)} \tag{6.18}$$

Suppose that we now take the network in Figure 6.3, consider the conductances G_1 and G_3 as linear and replace G_2 by a nonlinear function

$$i = g(v_{el})$$

where V_{el} is the voltage across this element,

$$v_{el} = v_1 - v_2$$

and g represents a nonlinear function. The two KCL equations are

$$G_1 v_1 + g(v_{el}) - J = 0$$
$$-g(v_{el}) + G_3 v_2 = 0$$

For the Newton–Raphson equation, we need the derivatives with respect to v_1 and v_2. Consider now only the nonlinear element. Using the chain rule of differentiation we can write

$$\frac{\partial g(v_{el})}{\partial v_1} = \frac{\partial g(v_{el})}{\partial v_{el}} \frac{\partial v_{el}}{\partial v_1} = +\frac{\partial g(v_{el})}{\partial v_{el}}$$

$$\frac{\partial g(v_{el})}{\partial v_2} = \frac{\partial g(v_{el})}{\partial v_{el}} \frac{\partial v_{el}}{\partial v_2} = -\frac{\partial g(v_{el})}{\partial v_{el}}$$

With these preliminary steps, we can now write the Newton–Raphson equation:

$$\begin{bmatrix} G_1 + \dfrac{\partial g(v_{el})}{\partial v_{el}} & -\dfrac{\partial g(v_{el})}{\partial v_{el}} \\ -\dfrac{\partial g(v_{el})}{\partial v_{el}} & G_3 + \dfrac{\partial g(v_{el})}{\partial v_{el}} \end{bmatrix} \begin{bmatrix} \Delta v_1 \\ \Delta v_2 \end{bmatrix} = -\begin{bmatrix} G_1 v_1 + g(v_{el}) - J \\ -g(v_{el}) + G_3 v_3 \end{bmatrix}$$

Comparing the Jacobian of the Newton–Raphson equation with the nodal formulation, we reach the important conclusion, valid for networks of any size:

1. Linear elements will be in the Jacobian in the same position as they were in the linear system matrix.
2. Nonlinear elements will have entries in the same positions as if they were linear; only their numerical values will be equal to the derivative $\partial g/\partial v_{el}$, evaluated with already available variables.

This conclusion will also be true for the other formulations, similar to the tableau or the modified nodal. So far, we considered only nonlinear resistive elements and the operating point.

Nonlinear storage elements (capacitors and inductors) contribute to the equations with their fluxes and charges. The current through the nonlinear capacitor is defined by

$$i_c = \frac{dq(v_c)}{dt} \tag{6.19}$$

where $q(v_c)$ is the charge and v_c is the voltage across the capacitor. The voltage across the nonlinear inductor is given by

$$v_L = \frac{d\phi(i_L)}{dt} \tag{6.20}$$

with ϕ denoting the flux.

Integration of systems with storage elements is always done by first replacing the derivative by a suitable algebraic expression and then solving the resulting nonlinear algebraic system by the Newton–Raphson method derived previously.

Many methods are available to replace the time domain derivatives by algebraic expressions. Books on numerical analysis usually describe the Runge–Kutta method. We mention it here because it is *not* suitable for solution of networks. There are several reasons for this, the main one being that the preferred modified nodal formulation does not lead to systems of differential, but rather to systems of algebraic-differential equations. Only two methods are widely used, the trapezoidal formula and a family of backward differentiation formulas (BDF). Among the BDFs, the simplest is the backward Euler, and we will base our explanations on this formula. It replaces the derivative by the difference of the previous and new value, divided by the step size, h,

$$\frac{dq(v_c)}{dt} \sim \frac{q_{new}(v_c) - q_{old}}{h}$$

$$\frac{d\phi(i_L)}{dt} \sim \frac{\phi_{new}(i_L) - \phi_{old}}{h}$$

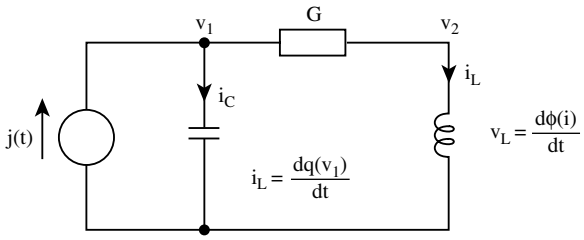


FIGURE 6.20 Network with a nonlinear capacitor and inductor.

Consider the network in Figure 6.20 with nonlinear storage elements and with a linear conductance, *G*. It can be described by three equations:

$$f_1 = \frac{dq(v_c)}{dt} + G(v_1 - v_2) - j(t) = 0$$

$$f_2 = -G(v_1 - v_2) + i_L = 0$$

$$f_3 = v_2 - \frac{d\phi(I_L)}{dt} = 0$$

The time derivatives are replaced by the backward Euler formula

$$\frac{q_{new}(v_1) - q_{old}}{h} + G(v_1 - v_2) - j(t) = 0$$

$$-G(v_1 - v_2) + i_L = 0$$

$$v_2 - \frac{\phi_{new}(i_L) - \phi_{old}}{h} = 0$$

thus changing the system into an algebraic one. If we now differentiate with respect to the variables v_1, v_2, and i_L, we obtain the Jacobian

$$\begin{bmatrix} \frac{1}{h}\frac{\partial q_{new}(v_1)}{\partial v_1} + G & -G & 0 \\ -G & +G & 1 \\ 0 & 1 & -\frac{1}{h}\frac{\partial \phi_{new}(i_L)}{\partial i} \end{bmatrix}$$

It can be observed that values of the derivatives are in the same places as would be the values of $C(L)$ of linear capacitors (inductors). In addition, the variable s from the Laplace domain is replaced by $1/h$.

The example used a grounded capacitor and a grounded inductor. Figure 6.21 gives the stamps for floating nonlinear elements and for the Newton–Raphson iteration, based on the backward Euler formula.

6.6 Nodal Analysis of Active Networks

Low-frequency analog filters are often built with active RC networks and the active elements are almost always operational amplifiers. We have seen in Section 6.4 that each such element adds one row and one

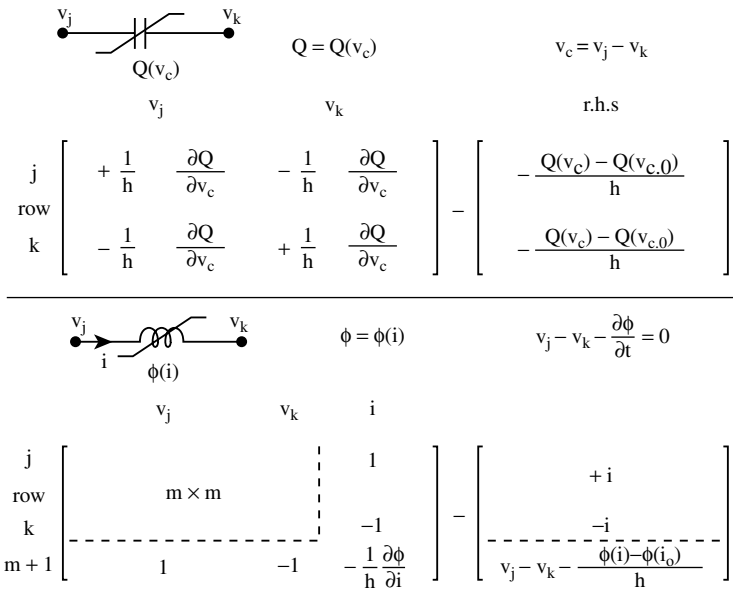


FIGURE 6.21 Stamp for a nonlinear capacitor and inductor.

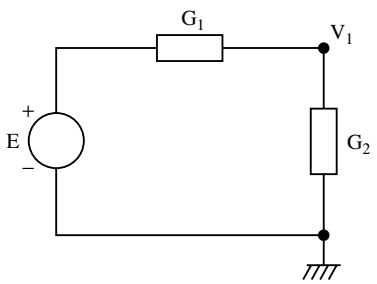


FIGURE 6.22 Example showing how to reduce the number of nodal equations.

column to the modified nodal system matrix, thus making the system too large for hand solutions. We need a method that can reduce the size of the matrix to the minimum. Such reduction is possible [1, 2], and becomes extremely simple if the voltage sources (dependent or independent) have one of their terminals grounded. Almost all practical networks meet this condition.

To introduce the method, consider the network in Figure 6.22. If we are not interested in the current through the voltage source, we can write only one nodal equation for the node on the right:

$$\left(G_1 + G_2\right) V_1 - EG_1 = 0$$

The source is known, the term in which it appears is transferred to the right side and, instead of three equations of the modified nodal formulation, we must solve only one. Consider next the network in Figure 6.23 with a voltage source and an ideal operational amplifier. One of the output terminals of the operational amplifier is grounded. Such amplifier is described by the equation

$$\left(V_+ - V_-\right)A = V_{out} \tag{6.21}$$

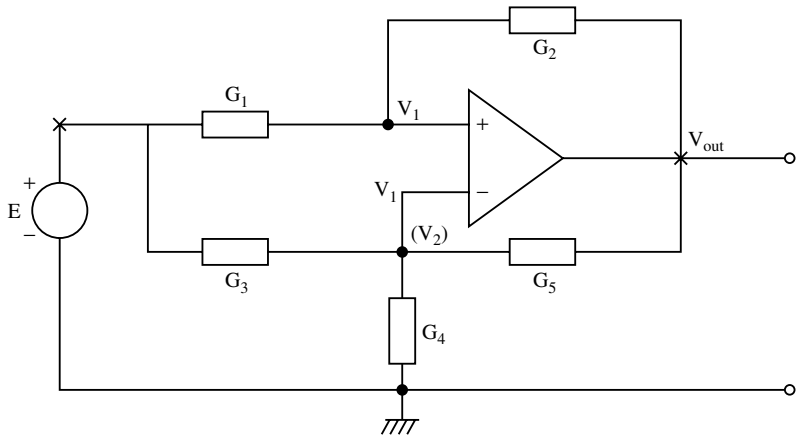


FIGURE 6.23 Nodal analysis of a network with one ideal operational amplifier.

where $A \to \infty$. Divide first by A and then substitute

$$B = -1/A \tag{6.22}$$

This changes the equation into

$$V_+ - V_- + BV_{out} = 0 \tag{6.23}$$

If the operational amplifier is ideal, set $B = 0$ and in such case $V_+ = V_-$; the operational amplifier will have the same voltages at its input terminals. We can take this into consideration, by simply writing the voltage with the same subscript to both input terminals of the operational amplifier, as was done in Figure 6.23. We are not interested in the current of the voltage source, nor in the current flowing into the operational amplifier. We mark our lack of interest by crossing out the nodes that have grounded voltage source; this was also done in Figure 6.23. Each node is given a voltage, but we write the nodal equations only at nodes that were not crossed out. For our example:

$$\left(G_1 + G_2\right) V_1 - G_2 V_{out} - EG_1 = 0$$

$$\left(G_3 + G_4 + G_5\right) V_1 - G_5 V_{out} - EG_3 = 0$$

Terms with the known source voltage are transferred to the right and we have the system

$$\begin{bmatrix} G_1 + G_2 & -G \\ G_3 + G_4 + G_5 & -G_5 \end{bmatrix} \begin{bmatrix} V_1 \\ V_{out} \end{bmatrix} = \begin{bmatrix} G_1 E \\ G_3 E \end{bmatrix}$$

A modified modal formulation would have required six equations.

This method can be used for any network if one node of each voltage source, dependent or independent, is grounded. All we have to do is assign every node a voltage, cross out nodes with voltage sources, and write nodal equations for the rest. It is advantageous to use conductances for resistors, because this way we avoid the fractions.

The method remains valid if the operational amplifier is not ideal and has the inverted gain *B*. The only difference is that for a nonideal amplifier we cannot make any assumptions on the voltages at its input terminals and the subscripts of such voltages must be different. This second case is also illustrated in Figure 6.23 by the voltage V_2 (in brackets) at the lower node. We still write nodal equations for the

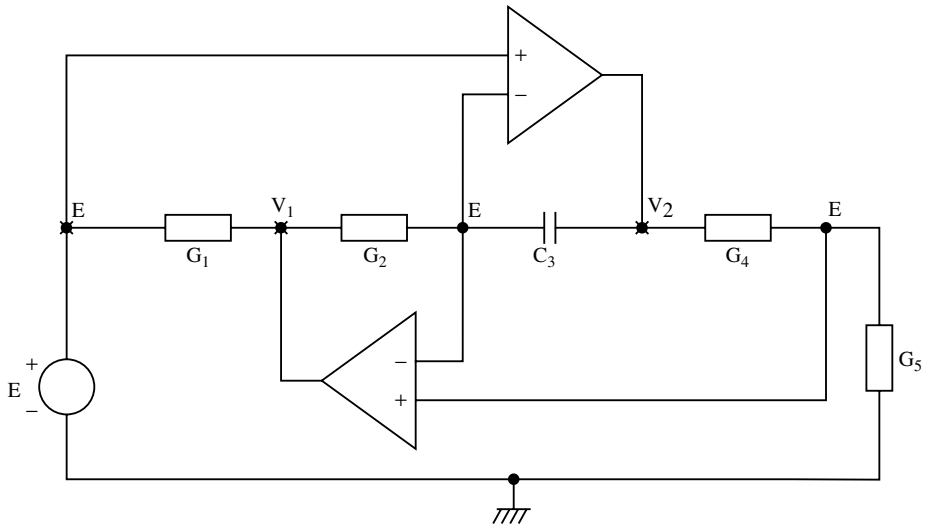


FIGURE 6.24 Nodal analysis of a network with two ideal operational amplifiers.

two input terminals, but to complete the system we must also attach the equation of the operational amplifier. The result is

$$V_1(G_1+G_2)-G_2V_{out}=G_1E$$
$$V_1(G_3+G_4+G_5)V_2-G_5V_{out}=G_3E$$
$$V_1-V_2+BV_{out}=0$$

and in matrix from

$$\begin{bmatrix} G_1+G_2 & 0 & -G_2 \\ 0 & G_3+G_4+G_5 & -G_5 \\ 1 & -1 & B \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \\ V_{out} \end{bmatrix}\begin{bmatrix} G_1E \\ G_3E \\ 0 \end{bmatrix}$$

where B can be set zero for an ideal operational amplifier.

We will give one example of a practical network, Figure 6.24. The operational amplifiers are ideal and thus the input voltage, E, appears at three terminals of the network. The other terminal voltages are marked by V_1 and V_2. The terminals with voltage sources are marked by crosses and only nodes 3 and 5, counting from left, remain for writing the KCL. They are

$$(G_2+sC_3)E-G_2V_1-sC_3V_2=0$$
$$(G_4+G_5)E-G_4V_2=0$$

Transferring terms containing the independent voltage source, E, to the other side of the equation, we arrive at the system

$$\begin{bmatrix} G_2 & sC_3 \\ 0 & G_4 \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix}=\begin{bmatrix} (G_2+sC_3)E \\ (G_4+G_5)E \end{bmatrix}$$

Conclusion

It was demonstrated that hand calculations should use nodal or mesh formulations. Computer applications should be based on modified nodal formulation. For active networks, it is advantageous to use the method of Section 6.6.

References

[1] J. Vlach, *Basic Network Theory with Computer Applications*, New York: Van Nostrand Reinhold, 1992.

[2] J. Vlach and K. Singhal, *Computer Methods for Circuit Analysis and Design*, 2nd ed., New York: Van Nostrand Reinhold, 1994.

7

Frequency Domain Methods

Peter Aronhime
University of Louisville

7.1 Network Functions

Network functions are employed to characterize linear, time-invariant networks in the zero state for a single excitation. Network functions contain information concerning a network's stability and natural modes. They allow a designer to focus on obtaining a desired output signal for a given input signal.

In this section, it is shown that the concept of network functions is obtained as an extension of the (transformed) element defining equations for resistors, capacitors, and inductors. The relationships of network functions to transformed loop and node equations are also described. As a result of these relationships, a list of properties of network functions can be generated, which is useful in the analysis of linear networks. Much is known about a network function for a given network even before an analysis is performed and the function itself is obtained.

Ohm's law, $v_R(t) = Ri_R(t)$ where R is in ohms, describes the relationship between the voltage across the resistor and the current through the resistor. These variables and their reference polarity and direction are depicted in Figure 7.1. If elements of the equation for Ohm's law are transformed, we obtain $V(s) = RI(s)$ because R is a constant. Thus, we obtain an Ohm's law-like expression in the frequency domain.

However, the capacitor's voltage and current are related by the integral

$$\begin{aligned} v_c(t) &= \frac{1}{C}\int_{-\infty}^{t} i_c(t)dt = \frac{1}{C}\int_{-\infty}^{0} i_C(t)dt + \frac{1}{C}\int_{0}^{t} i_C(t)dt \\ &= V_0 + \frac{1}{C}\int_{0}^{t} i_c(t)dt \end{aligned} \tag{7.1}$$

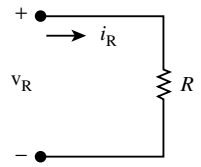


FIGURE 7.1 Reference polarity and direction for Ohm's law.

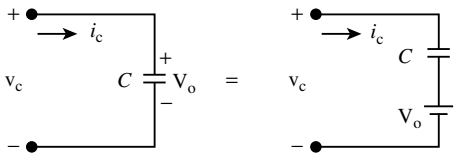


FIGURE 7.2 Capacitor representation showing reference polarities for voltages and reference direction for current.

where the voltage reference polarity and the current reference direction are illustrated in Figure 7.2 and the capacitance C is given in farads (F). Unlike the resistor, the relation between the capacitor voltage and the capacitor current is not a simple Ohm's law-like expression in the time domain. In addition, the voltage across the capacitor at any time t, is dependent on the entire history of the current through the capacitor.

The integral expression for the voltage across the capacitor can be split into two terms where the first term is the initial voltage across the capacitor, $V_0 = v_c(0)$. If the elements of the equation are transformed, we obtain

$$V_c(s) = \frac{V_0}{s} + \frac{1}{C}\frac{I_c(s)}{s} \tag{7.2}$$

Equation (7.2) shows that if $V_0 = 0$, then the expression for the transform of the capacitor voltage becomes more nearly Ohm's law-like in its form. Furthermore, if we associate the s that arises because of the integral of the current with the capacitor C, and define the *impedance* of the capacitor as $Z(s) = V_c(s)/I_c(s) = 1/(sC)$, then the equation becomes Ohm's law-like in the frequency domain.

A similar process can be applied to the inductor. The current through the inductor is expressed as

$$i_L(t) = \frac{1}{L}\int_{-\infty}^{t} v_L(t)dt = I_0 + \frac{1}{L}\int_{0}^{t} v_L(t)dt \tag{7.3}$$

where L is expressed in henries (H) and $I_0 = i_L(0)$ is the initial current through the inductor. Figure 7.3 depicts the reference polarity and direction for the inductor voltage and current. If the expression for the current through the inductor is transformed, the result is:

$$I_L(s) = \frac{I_o}{s} + \frac{1}{L}\frac{V_L(s)}{s} \tag{7.4}$$

Again, as with the capacitor, if $I_0 = 0$ and if the s that is included because of the integral of $v_L(t)$ is considered as associated with L, then the expression for the transform of the current through the inductor has an Ohm's law-like form if we define the *impedance* of the inductor as $Z(s) = V_L(s)/I_L(s) = sL$.

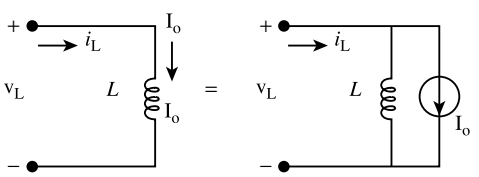


FIGURE 7.3 Inductor representation showing reference directions for currents and reference polarity for voltage.

The impedance concept is an important one in network analysis. It allows us to combine dissimilar elements in the frequency domain — something we cannot do in the time domain. In fact, impedance is a frequency domain concept. It is the ratio of the *transform* of the voltage across the port of the network to the *transform* of the current through the port with all independent sources within the network properly removed and with all initial voltages across capacitors and initial currents through inductors set to zero. Thus, when we indicate that independent sources are to be removed, we mean that initial conditions are to be set to zero as well.

The concept of impedance can be extended to linear, lumped, finite, time-invariant, one-port networks in general. We denote these networks as LLFT networks. These networks are linear. That is, they are composed of elements including resistors, capacitors, inductors, transformers, and dependent sources with parameters that are not functions of the voltage across the element or the current through the element. Thus, the differential equations describing these networks are linear.

These networks are lumped and not distributed. That is, LLFT networks do not contain transmission lines as network elements, and the differential equations describing these networks are ordinary and not partial differential equations.

LLFT networks are finite, meaning that they do not contain infinite networks and require only a finite number of network elements in their representation. Infinite networks are sometimes useful in modeling such things as ground connections in the surface of the earth, but we exclude the discussion of them here.

LLFT networks are time-invariant or constant instead of time-varying. Thus, the ordinary, linear differential equations describing LLFT networks have constant coefficients.

The steps for finding the impedance of an LLFT one-port network are:

1. Properly remove all independent sources in the network. By "properly" removing independent sources, we mean that voltage sources are replaced by short circuits and current sources are replaced by open circuits. Dependent sources are not removed.
2. Excite the network with a voltage source or a current source at the port, and find an equation or equations to solve for the other port variable.
3. Form $Z(s) = V(s)/I(s)$.

Simple networks do not need to be excited in order to determine their impedance, but in the general case an excitation is required. The next example illustrates these concepts.

Example 1. Find the impedances of the one-port networks in Figure 7.4.

Solution. The network in Figure 7.4(a) is composed of three elements connected in series. No independent sources are present, and there is zero initial voltage across the capacitor and zero initial current through the inductor. The impedance is determined as:

$$Z(s) = R + sL + \frac{1}{sC} = L\left(\frac{s^2 + s\frac{R}{L} + \frac{1}{LC}}{s}\right)$$

The network in Figure 7.4(b) includes a dependent source that depends on the voltage across R_1. The impedance of this network is not obvious, and so we should excite the port. Also, the capacitor has an

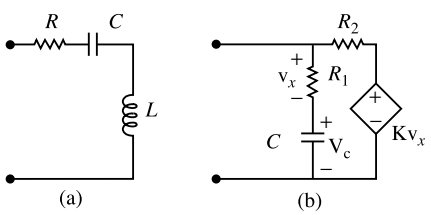


FIGURE 7.4 (a) A simple network. (b) A network containing a dependent source.

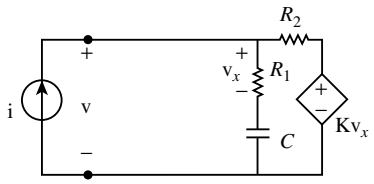


FIGURE 7.5 The network in Figure 7.4(b) prepared for analysis.

initial voltage V_c across it. This voltage is set to zero to find the impedance. Figure 7.5 is the network in Figure 7.4(b) prepared for finding the impedance. Using the impedance concept and two loop equations or two node equations, we obtain:

$$Z(s) = \frac{V(s)}{I(s)} = \frac{R_1 R_2 \left(s + \dfrac{1}{CR_1} \right)}{\left[R_1(1-K) + R_2 \right] \left(s + \dfrac{1}{C\left[R_1(1-K) + R_2 \right]} \right)}$$

□

The expressions for impedance found in the previous example are rational functions of s, and the coefficients of s are functions of the elements of the network including the coefficient K of the dependent source in the network in Figure 7.4(b). We will demonstrate that these observations are general for LLFT networks; but first we will extend the impedance concept in another direction.

We have defined the impedance of a one-port LLFT network. We can also define another network function — the admittance $Y(s)$. The admittance of a one-port LLFT network is the quotient of the *transform* of the current through the port to the *transform* of the voltage across the port with all independent sources within the network properly removed. One-port networks have only two linear network functions, impedance and admittance. Furthermore, $Z(s) = 1/Y(s)$ because both network functions concern the same port of the network, and the impedance or admittance relating the response to the excitation is the same whether a current excitation causes a voltage response or a voltage excitation causes a current response. An additional implication of this observation is that either network function can be determined with either type of excitation, voltage source or current source, applied to the network.

Figure 7.6 depicts a two-port network with the reference polarities and reference directions indicated for the port variables. Port one of the two-port network is formed from the two terminals labeled 1 and 1′. The two terminals labeled 2 and 2′ are associated to form port two. A two-port network has 12 network functions associated with it instead of only two, and so we will employ the following notation for these functions:

$$N_{RE}(s) = R(s)/E(s)$$

where $N_{RE}(s)$ is a network function, the subscript "R" is the port at which the response variable exists, the subscript "E" is the port at which the excitation is applied, $R(s)$ is the transform of the response variable, and $E(s)$ is the transform of the excitation that may be a current source or a voltage source depending on the particular network function. For example, for the two-port networks shown in Figure 7.7

$$Z_{21}(s) = \frac{V_2(s)}{I_1(s)} \text{ and } G_{12}(s) = \frac{V_1(s)}{V_2(s)} \tag{7.5}$$

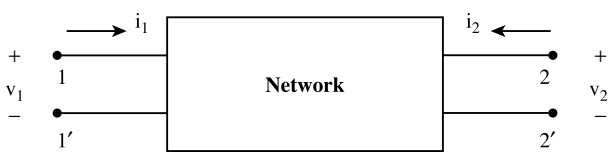


FIGURE 7.6 Reference polarities and reference directions for port variables of a two-port network.

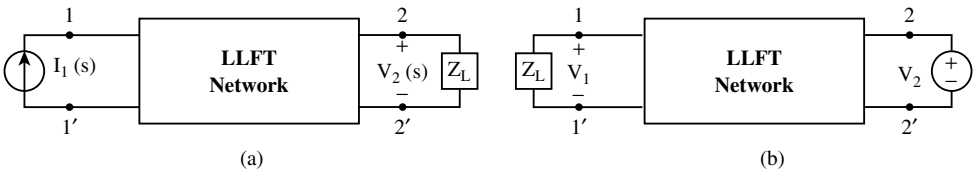


FIGURE 7.7 (a) Network configured for finding $Z_{21}(s)$. (b) Network for determining $G_{12}(s)$.

TABLE 7.1 Network Functions of Two-Port Networks

	Excitation Port	
Response Port	1	2
1	Z_{11}, Y_{11}	Z_{12}, Y_{12}, G_{12}, α_{12}
2	$Z_{21,}$ Y_{21}, $G_{21,}$ α_{21}	Z_{22}, Y_{22}

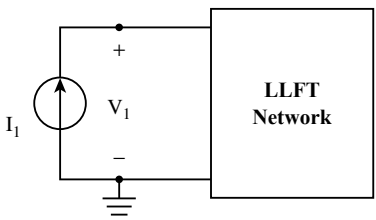


FIGURE 7.8 An LLFT network with n independent nodes plus the ground node.

Note that a load impedance has been placed across port two, the response port for $Z_{21}(s)$, in Figure 7.7(a). Also, a load has been connected across port one, the response port for $G_{12}(s) = V_1(s)/V_2(s)$ in Figure 7.7(b). It is assumed that all independent sources have been properly removed in both networks in Figure 7.7, and this assumption also applies to the loads. Of course, if a load impedance is changed, usually the network function will change. Thus, the load, if any, must be specified.

Table 7.1 lists the network functions of a two-port network. "G" denotes a voltage ratio, and "α" denotes a current ratio. The functions can also be grouped into driving-point and transfer functions. Driving-point functions are ones in which the excitation and response occur at the same port, and transfer network functions are ones in which the excitation and response occur at different ports. For example, $Z_{11}(s) = V_1(s)/I_1(s)$ is a driving-point network function, and $G_{21}(s) = V_2(s)/V_1(s)$ is a transfer network function. Of the twelve network functions for a two-port network, four are driving-point functions and eight are transfer functions. The two network functions for a one-port network are, of necessity, driving-point network functions.

Network functions are related to loop and node equations. Consider the LLFT network in Figure 7.8. Independent sources within the network have been properly removed. Assume the network has n independent nodes plus the ground node, and assume for simplicity that the network has no mutual inductance or dependent sources. Let us determine $Z_{11} = V_1/I_1$ in terms of a quotient of determinants of the nodal admittance matrix. The node equations, all written with currents leaving a node as positive currents, are

$$\begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} = \begin{bmatrix} I_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{7.6}$$

where V_i, $i = 1,2,\ldots, n$, are the unknown node voltages. The elements y_{ij}, $i, j = 1,2,\ldots, n$, of the nodal admittance matrix above have the form

$$y_{ij} = \pm\left\{g_{ij} + \frac{\Gamma_{ij}}{s} + sC_{ij}\right\} \tag{7.7}$$

where the plus sign is taken if $i = j$ and the minus sign is taken if i and j are unequal. The quantity g_{ij} is the sum of the conductances connected to node i if $i = j$, and if i does not equal j, it is the sum of the conductances connected between nodes i and j. A similar statement applies to C_{ij}. The quantity Γ_{ij} is the sum of the reciprocal inductances ($\Gamma = 1/L$) connected to node i if $i = j$, and it is the sum of the reciprocal inductances connected between nodes i and j if i does not equal j.

Solving for V_1 using Cramer's rule yields:

$$V_1 = \frac{\begin{vmatrix} I_1 & y_{12} & \cdots & y_{1n} \\ 0 & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ 0 & y_{n2} & \cdots & y_{nn} \end{vmatrix}}{\Delta'} \tag{7.8}$$

where Δ' is the determinant of the nodal admittance matrix. Thus,

$$V_1 = I_1 \frac{\Delta'_{11}}{\Delta'} \tag{7.9}$$

where Δ'_{11} is the cofactor of element y_{11} of the nodal admittance matrix. Thus, we can write

$$\frac{V_1}{I_1} = Z_{11} = \frac{\Delta'_{11}}{\Delta'} \tag{7.10}$$

If mutual inductance and dependent sources exist in the LLFT network, the nodal admittance matrix elements are modified. Furthermore, there may be more than one entry in the column matrix containing excitations. However, Z_{11} can still be expressed as a quotient of determinants of the nodal admittance matrix.

Next, consider the network in Figure 7.9, which is assumed to have n independent nodes plus a ground node. The response port exists between terminals j and k. In this network, we are making the pair of terminals j and k serve as the second port. Denote the **transimpedance** V_{jk}/I_1 as Z_{j1}. Let us express this transfer function Z_{j1} as a quotient of determinants. All independent sources within the network have been properly removed. Note that the node voltages are measured with respect to the ground terminal indicated, but the output voltage is the difference of the node voltages V_j and V_k. Thus, we have to solve

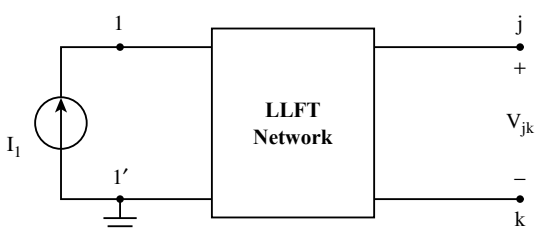


FIGURE 7.9 An LLFT network with two ports indicated.

for these two node voltages. Writing node equations, again taking currents leaving a node as positive currents, and solving for V_j using Cramer's rule we have

$$V_j = \frac{\begin{vmatrix} y_{11} & y_{12} & \cdots & y_{1(j-1)} & I_1 & y_{1k} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2(j-1)} & 0 & y_{2k} & \cdots & y_{2n} \\ \vdots & \vdots & \ddots & \vdots & \vdots & \vdots & \ddots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{n(j-1)} & 0 & y_{nk} & \cdots & y_{nn} \end{vmatrix}}{\Delta'} \tag{7.11}$$

which can be written as

$$V_j = I_1 \frac{\Delta'_{1j}}{\Delta'} \tag{7.12}$$

where Δ'_{1j} is the cofactor of element y_{1j} of the nodal admittance matrix. Similarly, we can solve for V_k and obtain

$$V_k = I_1 \frac{\Delta'_{1k}}{\Delta'} \tag{7.13}$$

Then, the transimpedance Z_{j1} can be expressed as

$$Z_{j1} = \frac{V_j - V_k}{I_1} = \frac{\Delta'_{1j} - \Delta'_{1k}}{\Delta'} \tag{7.14}$$

If terminal k in Figure 7.9 is common with the ground node so that the network is a grounded two-port network, then V_k is zero, and Z_{j1} can be expressed as Δ'_{1j}/Δ'. This result can be extended so that if the output voltage is taken between any node h and ground, then the transimpedance can be expressed as $Z_{h1} = \Delta'_{1h}/\Delta'$.

These results can be used to obtain an expression for G_{21} in terms of the determinants of the nodal admittance matrix. Figure 7.10 shows an LLFT network with a voltage excitation applied at port 1 and with port 2 open. The current $I_1(s)$ is given by V_1/Z_{11}. Then, V_2 is given by $V_2(s) = I_1 Z_{21}$. Thus,

$$G_{21} = \frac{V_2}{V_1} = \frac{I_1 Z_{21}}{I_1 Z_{11}} = \frac{\Delta'_{12}}{\Delta'_{11}} \tag{7.15}$$

Note that the determinants in the quotient are of equal order so that G_{21} is dimensionless.

Of course, network functions can also be expressed in terms of determinants of the loop impedance matrix. Consider the two-port network in Figure 7.11, which is excited with a voltage source applied to

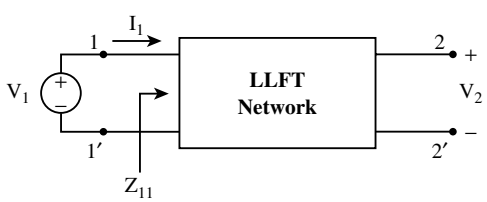


FIGURE 7.10 An LLFT network with port 2 open.

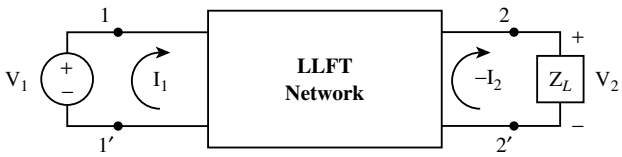


FIGURE 7.11 An LLFT network with load Z_L connected across port 2.

port 1 and has a load Z_L connected across port 2. Let us find the voltage transfer function G_{21} using loop equations. Assume that n independent loops exist, of which two are illustrated explicitly in Figure 7.11, and assume that no independent sources are present within the network. Also, assume for simplicity that the network contains no dependent sources or mutual inductance and that the loops are chosen so that V_1 is in only one loop. The loop equations are:

$$\begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nn} \end{bmatrix} \begin{bmatrix} I_1 \\ -I_2 \\ \vdots \\ I_n \end{bmatrix} = \begin{bmatrix} V_1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{7.16}$$

where I_j, $j = 1,3,\ldots,n$, and $-I_2$ are the loop currents, and the elements z_{ij} of the loop impedance matrix are given by:

$$z_{ij} = \pm\left(R_{ij} + sL_{ij} + \frac{D_{ij}}{s}\right), \quad i, j = 1, 2, ..., n \tag{7.17}$$

where we have assumed that all loop currents are taken in the same direction such as clockwise. The plus sign applies if $i = j$, and the minus sign is used if $i \neq j$. R_{ij} is the sum of the resistances in loop i if $i = j$, and R_{ij} is the sum of the resistances common to loops i and j if $i \neq j$. L_{ij} is the sum of the inductances in loop i if $i = j$, and it is the sum of the inductances common to loops i and j if $i \neq j$. A similar statement applies to the reciprocal capacitances $D_{ij}(D = 1/C)$. However, the element z_{22} includes the extra term Z_L which could be a quotient of determinants itself. Solving for $-I_2$ using Cramer's rule, we have Δ, which is the determinant of the $n \times n$ loop impedance matrix, and Δ_{12} is the cofactor of element z_{12} of the loop impedance matrix.

$$-I_2 = \frac{\begin{vmatrix} z_{11} & V_1 & z_{13} & \cdots & z_{1n} \\ z_{21} & 0 & z_{23} & \cdots & z_{2n} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ z_{n1} & 0 & z_{n3} & \cdots & z_{nn} \end{vmatrix}}{\Delta} = V_1 \frac{\Delta_{12}}{\Delta} \tag{7.18}$$

The transform voltage V_2 is given by $-I_2(s)Z_L$, and the transfer function G_{21} can be expressed as

$$G_{21} = \frac{V_2}{V_1} = Z_L \frac{\Delta_{12}}{\Delta} \tag{7.19}$$

Thus, G_{21} can be represented as a quotient of determinants of the loop impedance matrix multiplied by Z_L.

In a similar manner, we can write

$$Y_{jk} = \frac{\Delta_{kj}}{\Delta}, \quad j,k = 1,2 \tag{7.20}$$

Then, we can use this result to write:

$$V_1 = I_1 \frac{1}{Y_{11}} \text{ and } I_2 = V_1 Y_{21} \tag{7.21}$$

Thus,

$$\frac{I_2}{I_1} = \alpha_{21} = \frac{\Delta_{12}}{\Delta_{11}} \tag{7.22}$$

Table 7.2 summarizes some of these results.

TABLE 7.2 Network Function in Terms of Quotients of Determinants

$Y_{jk} = \Delta_{kj}/\Delta$	$G_{jk} = \Delta'_{kj}/\Delta'_{kk}$
$Z_{jk} = \Delta'_{kj}/\Delta'$	$\alpha_{jk} = \Delta_{kj}/\Delta_{kk}$

Properties of LLFT Network Functions

We now list properties of network functions of LLFT networks. These properties are useful as a check of an analysis of such networks.

1. Network functions of LLFT networks are real, rational functions of *s*, and therefore have the form

$$N(s) = \frac{P(s)}{Q(s)} = \frac{a_0 s^m + a_1 s^{m-1} + \cdots + a_{m-1} s + a_m}{b_0 s^n + b_1 s^{n-1} + \cdots + b_{n-1} s + b_n} \tag{7.23}$$

where the coefficients in both the numerator and denominator polynomials are real.

A network function of an LLFT network is a rational function because network functions can be expressed as quotients of determinants of nodal admittance matrices or of loop impedance matrices. The elements of these determinants are at most simple rational functions, and when the determinants are expanded and the fractions cleared, the result is always a rational function. The coefficients a_i, $i = 0,1,\ldots, m$, and b_j, $j = 0,1,\ldots, n$, are functions of the real elements of the network R, L, C, M, and coefficients of dependent sources. The constants R, L, C, and M are real. In most networks, the coefficients of dependent sources are real constants. Thus, the coefficients of LLFT network functions are real, and therefore the network function is a real function. It is possible for the "coefficients" of dependent sources in LLFT networks to themselves be real, rational functions of *s.* But when all the fractions are cleared, the result is a real, rational function.

2. An LLFT network function is completely defined by its self-poles, self-zeros, and the scale factor H.

If the numerator and denominator polynomials are factored and there are no common factors, we have

$$N(s) = \frac{P(s)}{Q(s)} = \frac{a_0}{b_0} \frac{(s-z_1)(s-z_2)\cdots(s-z_m)}{(s-p_1)(s-p_2)\cdots(s-p_n)} \tag{7.24}$$

where $a_0/b_0 = H$ is the scale factor. The values of $s = z_1, z_2,\ldots, z_m$ are zeros of the polynomial $P(s)$ and self-zeros of the network function. Also, $s = p_1, p_2,\ldots,p_n$ are zeros of the polynomial $Q(s)$ and self-poles

of the network function. In addition, $N(s)$ may have other poles or zeros at infinity. We call these poles and zeros **mutual poles** and **mutual** zeroes because they result from the difference in degrees of the numerator and denominator polynomials.

3. Counting both self and mutual poles and zeros, and counting a k^{th} order pole or zero k times, N(s) has the same number of poles as zeros, and this number is equal to the highest power of s in $N(s)$.

If $m = n$, then there are n self-zeros and n self-poles and no mutual poles or zeros. If $n > m$, then there are m self-zeros and n self-poles. There are also $n - m$ mutual zeros. Thus, there are n poles and $m + (n - m) = n$ zeros. A similar statement can be constructed for $n < m$.

4. Complex roots of $P(s)$ and $Q(s)$ occur in conjugate pairs.

This property follows from the fact that the coefficients of the numerator and denominator polynomials are real. Thus, complex factors of these polynomials have the form

$$(s+c+jd)(s+c-jd)=\left[(s+c)^2+d^2\right] \tag{7.25}$$

where c and d are real constants.

5. A driving point function of a network having no dependent sources can have neither poles nor zeros in the right-half s-plane (RHP), and poles and zeros on the imaginary axis must be simple. The same restrictions apply to the poles of transfer network functions of such networks but not to the zeros of transfer network functions.

Elsewhere in this handbook it is shown that the denominator polynomials of LLFT networks having no dependent sources cannot have RHP roots, and roots on the imaginary axis, if any, must be simple. However, the reciprocal of a driving-point network function is also a network function. For example, $1/Y_{22} = Z_{22}$. Thus, restrictions on the locations of poles of driving-point network functions also apply to zeros of driving-point network functions.

However, the reciprocal of a transfer network function is not a network function (see [5]). For example, $1/Y_{21} \neq Z_{21}$. Thus, restrictions on the poles of a transfer function do not apply to its zeros.

We can make a classification of the factors corresponding to the allowed types of poles as follows:

TABLE 7.3 A Classification of Factors of Network Functions of *LLFT* Networks Containing No Dependent Sources

Type	Factor(s)	Conditions
A	$(s + a)$	$a \geq 0$
B	$(s + b + jc)(s + b - jc)$	$b > 0, c > 0$
C	$(s + jd)(s - jd)$	$d > 0$

The Type A factor corresponds to a pole on the $-\sigma$ axis. If $a = 0$, then the factor corresponds to a pole on the imaginary axis, and so only one such factor is allowed. Type B factors correspond to poles in the left-half s-plane (LHP), and Type C factors correspond to poles on the imaginary axis.

6. The coefficients of the numerator and denominator polynomials of a driving-point network function of an LLFT network with no dependent sources are positive. The coefficients of the denominator polynomial of a transfer network function are all one sign. Without loss of generality, we take the sign to be positive. But some or all of the coefficients of the numerator polynomial of a transfer network function may be negative.

A polynomial made up of the factors listed in Table 7.3 would have the form:

$$Q(s)=(s+a_1)\cdots\left[(s+b_1)^2+c_1^2\right]\cdots(s^2+d_1^2)\cdots$$

Note that all the constants are positive in the expression for $Q(s)$, and so it is impossible for any of the coefficients of $Q(s)$ to be negative.

7. There are no missing powers of *s* in the numerator and denominator polynomials of a driving-point network function of an LLFT network with no dependent sources unless all even or all odd powers of *s* are missing or the constant term is missing. This statement holds for the denominator polynomials of transfer functions of such networks, but there may be missing powers of *s* in the numerator polynomials of transfer functions.

Property 7 is easily illustrated by combining types of factors from Table 7.3. Thus, a polynomial consisting only of type A factors contains all powers of *s* between the highest power and the constant term unless one of the "*a*" constants is zero. Then, the constant term is missing. Two *a* constants cannot be zero because then there would be two roots on the imaginary axis at the same location. The roots on the imaginary axis would not be simple.

A polynomial made up of only type B factors contains all powers of *s*, and a polynomial containing only type C factors contains only even powers of *s*. A polynomial constructed from type C factors except for one A type factor with $a = 0$ contains only odd powers of *s*. If a polynomial is constructed from type B and C factors, then it contains all power of *s*.

8. The orders of the numerator and denominator polynomials of a driving-point network function of an LLFT network, which contains no dependent sources can differ by no more than one.

The limiting behavior at high frequency must be that of an inductor, a resistor, or a capacitor. That is, if $N_{dp}(s)$ is a driving-point network function, then

$$\lim_{s\to\infty} N_{dp}(s) = \begin{cases} K_1 s \\ K_2 \\ K_3/s \end{cases}$$

where K_i, = 1, 2, 3, are real constants.

9. The terms of lowest order in the numerator and denominator polynomials of a driving-point network function of an LLFT network containing no dependent sources can differ in order by no more than one.

The limiting behavior at low frequency must be that of an inductor, a resistor, or a capacitor. That is,

$$\lim_{s\to 0} N_{dp}(s) = \begin{cases} K_4 s \\ K_5 \\ K_6/s \end{cases}$$

where the constants K_i, $i = 4, 5, 6$, are real constants.

10. The maximum order of the numerator polynomials of the dimensionless transfer functions G_{12}, G_{21}, α_{12}, and α_{21}, of an LLFT network containing no dependent sources is equal to the order of the denominator polynomials. The maximum order of the numerator polynomial of the transfer functions Y_{12}, Y_{21}, Z_{12}, and Z_{21} is equal to the order of the denominator polynomial plus 1. However, the minimum order of the numerator polynomial of any transfer function may be zero.

If dependent sources are included in an LLFT network, then it is *possible* for the network to have poles in the RHP or multiple poles at locations on the imaginary axis. However, an important application of stable networks containing dependent sources is to mimic (simulate) the behavior of LLFT networks that contain no dependent sources. For example, networks that contain resistors, capacitors, and dependent sources can mimic the behavior of networks containing only resistors, capacitors, and inductors. Thus, low-frequency filters can be constructed without the need for heavy, expensive inductors that would ordinarily be required in such applications.

7.2 Network Theorems

In this section, we provide techniques, strategies, equivalences, and theorems for simplifying the analysis of LLFT networks or for checking the results of an analysis. They can save much work in the analysis of some networks if one remembers to apply them. Thus, it is convenient to have them listed in one place. To begin, we list nine equivalences that are often called source transformations.

Source Transformations

Table 7.4 is a collection of memory aids for the nine source transformations. Source transformations are simple ways the elements and sources externally connected to a network N can be combined or eliminated without changing the voltages and currents within network N thereby simplifying the problem of finding a voltage or current within N.

Source transformation one in Table 7.4 shows the equivalence between two voltage sources connected in series and a single voltage source having a value that is the sum of the voltages of the two sources. A double-headed arrow is shown between the two network representations because it is sometimes advantageous to use this source transformation in reverse. For example, if a voltage source that has both DC and AC components is applied to a linear network N, it may be useful to represent that voltage source as two voltage sources in series — one a DC source and the other an AC source.

Source transformation two shows two voltage sources connected in parallel. Unless V_1 and V_2 are equal, the network would not obey Kirchhoff's law as evidenced by a loop equation written in the loop formed by the two voltage sources. A network that does not obey Kirchhoff's laws is termed a contradiction. Thus, a single-headed arrow is shown between the two network representations.

Source transformations three and four are duals, respectively, of source transformations two and one. The current sources must be equal in transformation three or else Kirchhoff's law would not be valid at the node indicated, and the circuit would be a contradiction.

Source transformation five shows that the circuit M_1 can be removed without altering any of the voltages and currents inside *N*. Whether M_1 is connected as shown or is removed, the voltage applied to N remains V_s. However, the current supplied by the source V_s changes from I_s to I_1.

Source transformation six shows that circuit M_2 can be replaced by a short circuit without affecting voltages and currents in *N*. Whether M_2 is in series with the current source I_1 as shown or removed, the current applied to N is the same. However, if network M_2 is removed, then the voltage across the current source changes from V_s to V_1.

Source transformation seven is sometimes termed a Thévenin circuit to Norton circuit transformation. This transformation, as shown by the double-headed arrow, can be used in the reverse direction. Thévenin's theorem is discussed thoroughly later in this section.

Source transformation eight is sometimes described as "pushing a voltage source through a node," but we will term it as "splitting a voltage source." Loop equations remain the same with this transformation, and the current leaving network N through the lowest wire continues to be I_s.

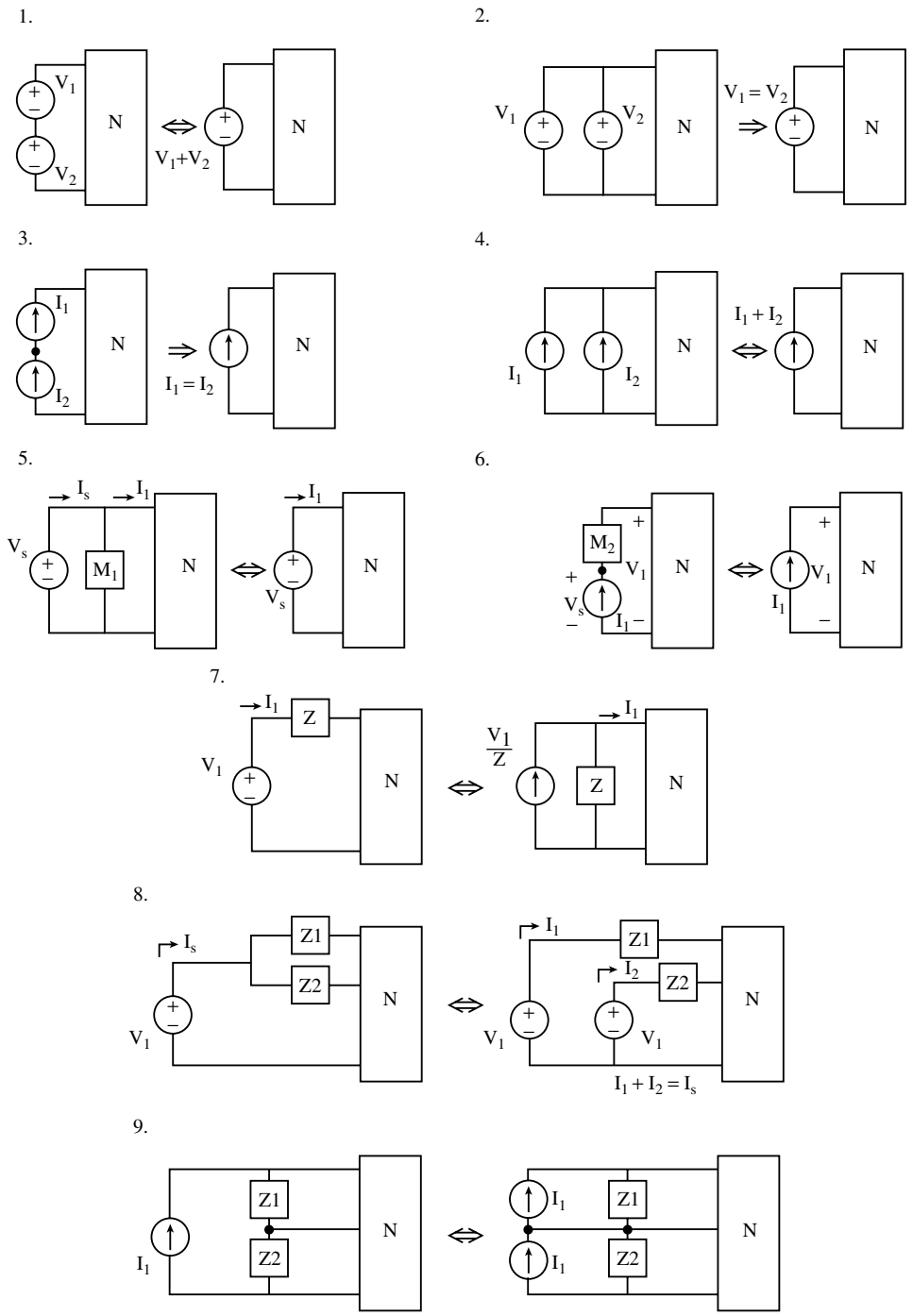
Source transformation nine shows that if a current source is not in parallel with one element, then it can be "split" as shown. Now, each one of the current sources I_1 has an impedance in parallel. Thus, analysis of network N may be simplified because source transformation seven can be applied.

Source transformations cannot be applied to all networks, but when they can be employed, they usually yield useful simplifications of the network.

Example 2. Use source transformations to find V_0 for the network shown. Initial current through the inductor in the network is zero.

Solution. The network can be readily simplified by employing source transformation five from Table 7.4 to eliminate R_1 and also I_2. Then, source transformation six can be used to eliminate V_1 because it is an element in series with a current source. The results to this point are illustrated in Figure 7.13(a). If we

TABLE 7.4 Source Transformations



N is an arbitrary network in which analysis for a voltage or current is to be performed. M_1 is an arbitrary one-port network or network element except a voltage source. M_2 is an arbitrary one-port network or network element except a current source. It is assumed there is no magnetic coupling between *N* and M_1 or M_2. There are no dependent sources in *N* in Source Transformation 5 that depend on I_s. Furthermore, there are no dependent sources in *N* in Source Transformation 6 that depend on V_s. However, M_1 and M_2 can have dependent sources that depend on voltages or currents in *N*. *Z*, Z_1 and Z_2 are one-port impedances.

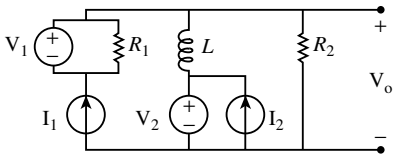


FIGURE 7.12 Network for Example 2.

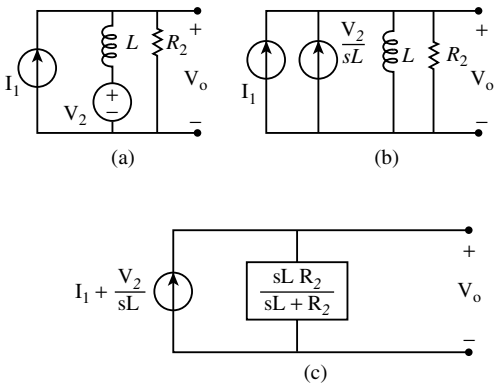


FIGURE 7.13 (a, b, c) Applications of source transformations to the network in Figure 7.12.

then apply transformation seven, we obtain the network in Figure 7.13(b). Next, we can apply transformation four to obtain the single loop network in Figure 7.13(c). The output voltage can be written in the frequency domain as

$$V_0 = \left(I_1 + \frac{V_2}{sL}\right)\left(\frac{sLR_2}{sL + R_2}\right)$$

Source transformations can often be used advantageously with the following theorems. □

Dividers

Current dividers and voltage dividers are circuits that are employed frequently, especially in the design of electronic circuits. Thus, dividers must be analyzed quickly. The relationships derived next satisfy this need.

Figure 7.14 is a current divider circuit. The source current I_s divides between the two impedances, and we wish to determine the current through Z_2. Writing a loop equation for the loop indicated, we have

$$I_2 Z_2 - (I_s - I_2)Z_1 = 0 \tag{7.26}$$

from which we obtain

$$I_2 = I_s \frac{Z_1}{Z_1 + Z_2} \tag{7.27}$$

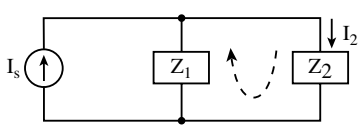


FIGURE 7.14 A current divider.

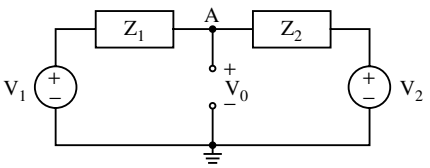


FIGURE 7.15 Enhanced voltage divider.

A circuit that we term an **enhanced** voltage divider is depicted in Figure 7.15. This circuit contains two voltage sources instead of the usual single source, but the enhanced voltage divider occurs more often in electronic circuits. Writing a node equation at node A and solving for V_0, we obtain

$$V_0 = \frac{V_1 Z_2 + V_2 Z_1}{Z_1 + Z_2} \tag{7.28}$$

If V_2, for example, is zero, then the results from the enhanced voltage divider reduce to those of the single source voltage divider.

Example 3. Use (7.28) to find V_0 for the network in Figure 7.16.

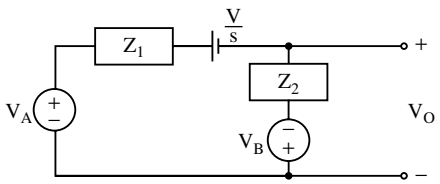


FIGURE 7.16 Circuit for Example 3.

Solution. The network in Figure 7.16 matches with the network used to derive (7.28) even though it is drawn somewhat differently and has three voltage sources instead of two. However, we can use (7.28) to write the answer for V_0 by inspection.

$$V_0 = \frac{\big(V_A - (V/s)\big)Z_z - V_B Z_1}{Z_1 + Z_2}$$

□

The following example illustrates the use of source transformations together with the voltage divider.

Example 4. Find V_0 for the network shown in Figure 7.17. The units of K, the coefficient of the dependent source, are ohms, and the capacitor is initially uncharged.

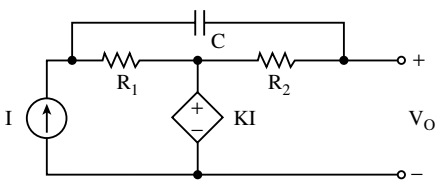


FIGURE 7.17 Network for Example 4.

Solution. We note that the dependent voltage source is not in series with any one particular element and that the independent current source is not in parallel with any one particular element. However, we can split both the voltage source and the current source using source transformations eight and nine, respectively, from Table 7.4. Then, employing transformations five and seven, we obtain the network configuration depicted in Figure 7.18, for which we can use the voltage divider to write:

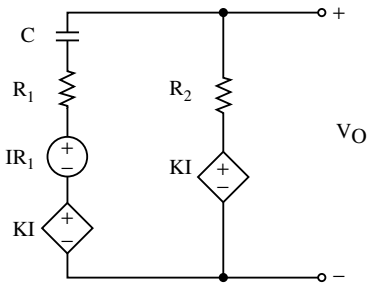


FIGURE 7.18 Results after employing source transformations on the network in Figure 7.17.

$$V_0 = \frac{I(K + R_1)R_2 + KI(R_1 + (1/sC))}{R_1 + R_2 + (1/sC)}$$

It should be mentioned that the method used to find V_0 in this example is not the most efficient one. For example, loops can be chosen for the network in Figure 7.17, so only one unknown loop is current. However, source transformations and dividers become more powerful analysis tools as they are coupled with additional network theorems.

Superposition

Superposition is a property of all linear networks, and whether it is used directly in the analysis of a network or not, it is a concept that is valuable in thinking about LLFT networks. Consider the LLFT network shown in Figure 7.19 in which, say, we wish to solve for I_1. Assume the network has n independent loops, and, for simplicity, assume no sources are within the box in the figure and that initial voltages across capacitors and initial currents through inductors are zero or are represented by independent sources external to the box. Note that one dependent source is shown in Figure 7.19 that depends on a voltage V_x in the network and that two independent sources, V_1 and V_2, are applied to the network. If loops are chosen so that each source has only one loop current flowing through it as indicated in Figure 7.19, then the loop equations can be written as

$$\begin{bmatrix} V_1 \\ V_2 \\ KV_x \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \cdots & \cdots & \ddots & \cdots \\ z_{n1} & z_{n2} & \cdots & z_{nn} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} \tag{7.29}$$

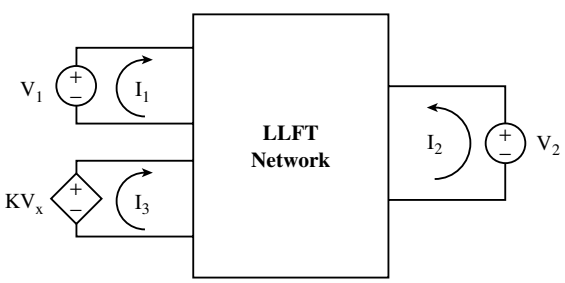


FIGURE 7.19 LLFT network with three voltage sources, of which one is dependent.

where the elements of the loop impedance matrix are defined in the section describing network functions. Solving for I_1 using Cramer's rule, we have:

$$I_1 = V_1 \frac{\Delta_{11}}{\Delta} + V_2 \frac{\Delta_{21}}{\Delta} + KV_x \frac{\Delta_{31}}{\Delta} \tag{7.30}$$

where Δ is the determinant of the loop impedance matrix, and Δ_{j1}, $j = 1, 2, 3$, are cofactors. The expression for I_1 given in (7.30) is an intermediate and not a finished solution. The finished solution would express I_1 in terms of the independent sources and the parameters (Rs, Ls, Cs, Ms, and Ks) of the network and not in terms of an unknown V_x. Thus, one normally has to eliminate V_x from the expression for I_1; but the intermediate expression for I_1 illustrates superposition. There are three components that add up to I_1 in (7.30) — one for each source including one for the dependent source. Furthermore, we see that each source is multiplied by a transadmittance (or a driving-point admittance in the case of V_1). Thus, we can write:

$$I_1 = V_1 Y_{11} + V_2 Y_{12} + KV_x Y_{13} \tag{7.31}$$

where each admittance is found from the port at which a voltage source (whether independent or dependent) is applied. The response variable for each of these admittances is I_1 at port 1.

The simple derivation that led to (7.30) is easily extended to both types of independent excitations (voltage sources and current sources) and to all four types of dependent sources. The generalization of (7.30) leads to the conclusion:

> To apply superposition in the analysis of a network containing at least one independent source and a variety of other sources, dependent or independent, one finds the contribution to the response from each source in turn with all other source, dependent or independent, properly removed and then adds the individual contributions to obtain the total response. No distinction is made between independent and dependent sources in the application of superposition other than requiring the network to have at least one independent source.

However, if dependent sources are present in the network, the quantities (call them V_x and I_x) on which the dependent sources depend must often be eliminated from the answer by additional analysis if the answer is to be useful unless V_x or I_x are themselves the variables of independent sources or the quantities sought in the analysis.

Some examples will illustrate the procedure.

Example 5. Find V_0 for the circuit shown using superposition. In this circuit, only independent sources are present.

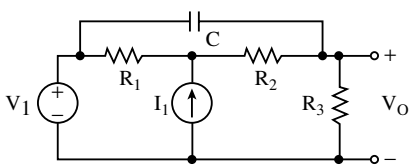


FIGURE 7.20 Network for Example 5.

Solution. Two sources in the network, therefore, we abstract two fictitious networks from Figure 7.20. The first is shown in Figure 7.21(a) and is obtained by properly removing the current source I_1 from the original network. The impedance of the capacitor can then be combined in parallel with $R_1 + R_2$, and the contribution to V_0 from V_1 can be found using a voltage divider. The result is

$$V_0 \text{ due to } V_1 = V_1 \frac{s + \dfrac{1}{C(R_1 + R_2)}}{s + \dfrac{R_1 + R_2 + R_3}{C(R_1 + R_2)R_3}}$$

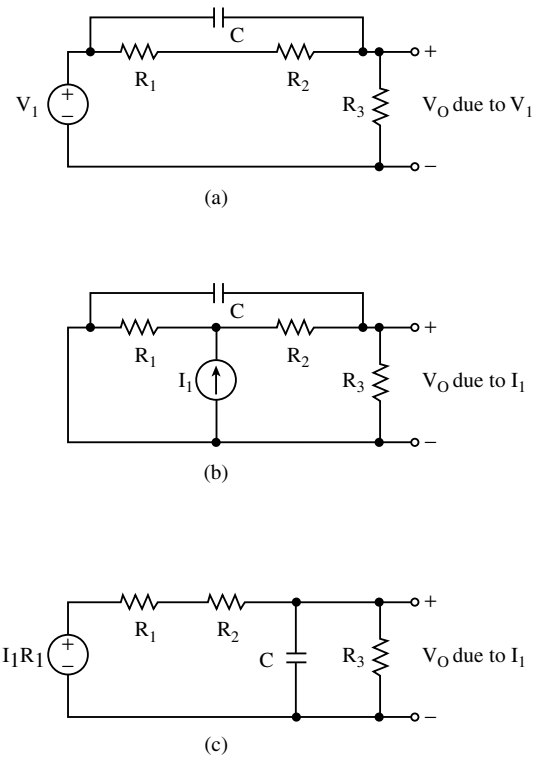


FIGURE 7.21 (a, b, c) Steps in the use of superposition for finding the response to two independent sources.

The second fictitious network, shown in Figure 7.21(b), is obtained form the original network by properly removing the voltage source V_1. Redrawing the circuit and employing source transformation seven (in reverse) yields the circuit in Figure 7.21(c). Again, employing a voltage divider, we have

$$V_0 \text{ due to } I_1 = I_1 \frac{\dfrac{R_1}{C(R_1+R_2)}}{s+\dfrac{R_1+R_2+R_3}{C(R_1+R_2)R_3}}$$

Then, adding the two contributions, we obtain

$$V_0 = \frac{V_1\left[s+\dfrac{1}{C(R_1+R_2)}\right]+I_1\dfrac{R_1}{C(R_1+R_2)}}{s+\dfrac{R_1+R_2+R_3}{C(R_1+R_2)R_3}}$$

The next example includes a dependent source.

Example 6. Find i in the network shown in Figure 7.22 using superposition.

Solution. Since there are two sources, we abstract two fictitious networks from Figure 7.22. The first one is shown in Figure 7.23(a) and is obtained by properly removing the dependent current source. Thus,

$$i \text{ due to } v_1 = \frac{v_1}{R_1+R_2}$$

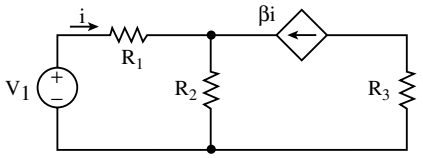


FIGURE 7.22 Network for Example 6.

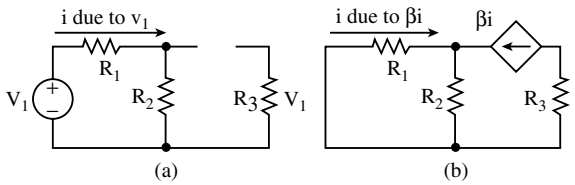


FIGURE 7.23 (a, b) Steps in the application of superposition to the network in Figure 7.22.

Next, voltage source v_1 is properly removed yielding the fictitious network in Figure 7.23(b). An important question immediately arises about this network. Namely, why is not i in this network zero? The reason i is not zero is that the network in Figure 7.23(b) is merely an abstracted network that concerns a step in the analysis of the original circuit. It is an artifice in the application of superposition, and the dependent source is considered to be independent for this step. Thus,

$$i \text{ due to } \beta i = -\frac{\beta i R_2}{R_1 + R_2}$$

Adding the two contributions, we obtain the intermediate result:

$$i = \frac{v_1}{R_1 + R_2} - \frac{\beta i R_2}{R_1 + R_2}$$

Collecting the terms containing i, we obtain the finished solution for i:

$$i = \frac{v_1}{(\beta + 1)R_2 + R_1}$$

We note that the finished solution depends only on the independent source v_1 and parameters of the network, which are R_1, R_2, and β. □

The following example involves a network in which a dependent source depends on a voltage that is neither the voltage of an independent source nor the voltage being sought in the analysis.

Example 7. Find V_0 using superposition for the network shown in Figure 7.24. Note that K, the coefficient of the VCCS, has the units of siemens.

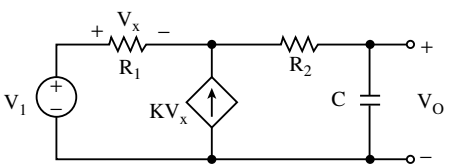


FIGURE 7.24 Network for Example 7.

Solution. When the dependent current source is properly removed, the network reduces to a simple voltage divider, and the contribution to V_0 due to V_1 can be written as:

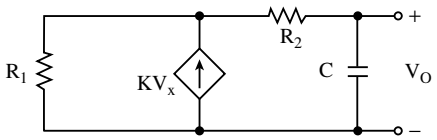


FIGURE 7.25 Fictitious network obtained when the voltage source is properly removed in Figure 7.24.

$$V_0 \text{ due to } V_1 = V_1 \frac{1}{sC(R_1 + R_2) + 1}$$

Then, reinserting the current source and properly removing the voltage source, we obtain the fictitious network shown in Figure 7.25. Using the current divider to obtain the current flowing through the capacitor and then multiplying this current by the impedance of the capacitor, we have:

$$V_0 \text{ due to } KV_x = \frac{KV_xR_1}{sC(R_1 + R_2) + 1}$$

Adding the individual contributions to form V_0 provides the equation

$$V_0 = \frac{V_1 + KV_xR_1}{sC(R_1 + R_2) + 1}$$

This is a valid expression for V_0. It is not a finished expression however, because it includes V_x, an unknown voltage. Superposition has taken us to this point in the analysis, but more work must be done to eliminate V_x. However, superposition can be applied again to solve for V_x, or other analysis tools can be used. The results for V_x are:

$$V_x = \frac{V_1 sCR_1}{sC[R_1 + R_2 + R_1R_2K] + R_1K + 1}$$

Then, eliminating V_x from the equation for V_0, we obtain the finished solution as:

$$V_0 = V_1 \frac{R_1K + 1}{sC[R_1 + R_2 + R_1R_2K] + R_1K + 1}$$

Clearly, superposition is not the most efficient technique to use to analyze the network in Figure 7.24. Analysis based on a node equation written at the top end of the current source would yield a finished result for V_0 with less algebra. However, this example does illustrate the application of superposition when a dependent source depends on a rather arbitrary voltage in the network. □

If the dependent current source in the previous example depended on V_1 instead of on the voltage across R_1, the network would be a different network. This is illustrated by the next example.

Example 8. Use superposition to determine V_0 for the circuit in Figure 7.26.

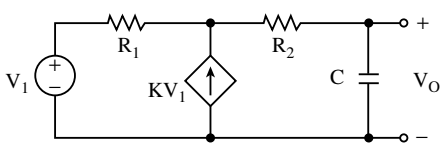


FIGURE 7.26 Network for Example 8.

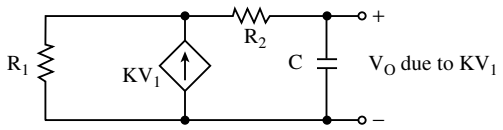


FIGURE 7.27 A step in the application of superposition to the network in Figure 7.26.

Solution. If the current source is properly removed, the results are the same as for the previous example. Thus,

$$V_0 \text{ due to } V_1 = \frac{V_1}{sC(R_1 + R_2) + 1}$$

Then, if the current source is reinserted, and the voltage source is properly removed, we have the circuit depicted in Figure 7.27. A question that can be asked for this circuit is why include the dependent source KV_1 if the voltage on which it depends, namely V_1, has been set to zero? However, the network shown in Figure 7.27 is merely a fictitious network that serves as an aid in the application of superposition, and superposition deals with all sources, whether they are dependent or independent, as if they were independent. Thus, we can write:

$$V_0 \text{ due to } KV_1 = \frac{KV_1R_1}{sC(R_1 + R_2) + 1}$$

Adding the contributions to form V_0, we obtain

$$V_0 = V_1 \frac{KR_1 + 1}{sC(R_1 + R_2) + 1}$$

and this is the finished solution.

In this example, we did not have the task of eliminating an unknown quantity from an intermediate result for V_0 because the dependent source depended on an independent source V_1, which is assumed to be known. □

Superposition is often useful in the analysis of circuits having only independent sources, but it is especially useful in the analysis of some circuits having both independent and dependent sources because it deals with all sources as if they were independent.

Thévenin's Theorem

Thévenin's theorem is useful in reducing the complexity of a network so that analysis of the network for a particular voltage or current can be performed more easily. For example, consider Figure 7.28(a), which is composed of two subnetworks A and B that have only two nodes in common. In order to facilitate analysis in subnetwork B, it is convenient to reduce subnetwork A to the network in Figure 7.28(b) which is termed the Thévenin equivalent of subnetwork A. The requirement on the Thévenin's equivalent

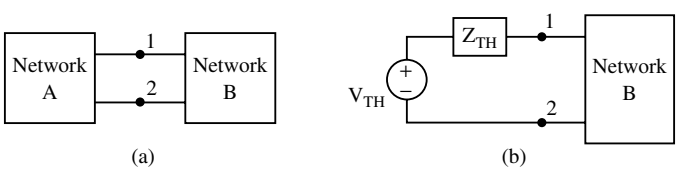


FIGURE 7.28 (a) Two subnetworks having a common pair of terminals. (b) The Thévenin equivalent for subnetwork A.

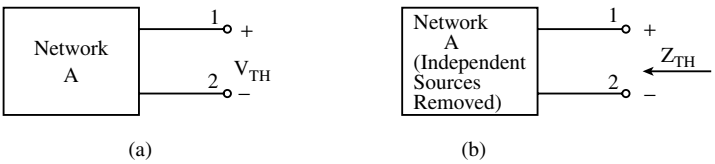


FIGURE 7.29 (a) Network used for finding V_{TH}. (b) Network used for obtaining $Z_{TH.}$

network is that, when it replaces subnetwork A, the voltages and currents in subnetwork B remain unchanged. We assume that no inductive coupling occurs between the subnetworks, and that dependent sources in B are not dependent on voltages or currents in A. We also assume that subnetwork A is an LLFT network, but subnetwork B does not have to meet this assumption.

To find the Thévenin equivalent network, we need only determine V_{TH} and Z_{TH}. V_{TH} is found by unhooking B from A and finding the voltage that appears across the terminals of A. In other words, we abstract a fictitious network from the complete network as depicted in Figure 7.29(a), and find the voltage that appears between the terminals that were common to B. This voltage is V_{TH}.

Z_{TH} is also obtained from a fictitious network that is created from the fictitious network used for finding V_{TH} by properly removing all independent sources. The effects that dependent sources have on the procedure are discussed later in this section. The fictitious network used for finding Z_{TH} is depicted in Figure 7.29(b). Oftentimes, the expression for Z_{TH} cannot be found by mere inspection of this network, and, therefore, we must excite the network in Figure 7.29(b) by a voltage source or a current source and find an expression for the other variable at the port in order to find Z_{TH}.

Example 9. Find the Thévenin equivalent of subnetwork A in Figure 7.30.

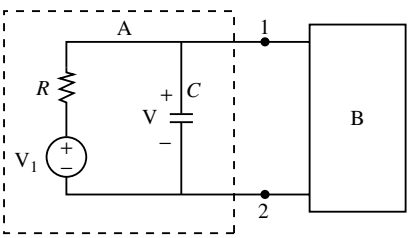


FIGURE 7.30 Network for Example 9.

Solution. No dependent sources exist in subnetwork A, but the capacitor has an initial voltage V across it. However, the charged capacitor can be represented by an uncharged capacitor in series with a transformed voltage source V/s. The fictitious network used for finding V_{TH} is given in Figure 7.31(a).

It should be noted that although subnetwork B has been removed and the two terminals that were connected to B are now "open circuited" in Figure 7.31(a), current is still flowing in network A. V_{TH} is easily obtained using a voltage divider:

$$V_{TH} = \frac{V_1(s) + VCR}{sCR + 1}$$

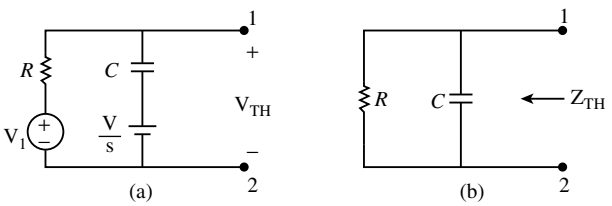


FIGURE 7.31 (a) Network of finding V_{TH}. (b) Network that yields Z_{TH}.

Z_{TH} is obtained from the fictitious network in Figure 7.31(b), which is obtained by properly removing the independent source and the voltage representing the initial voltage across the capacitor in Figure 7.31(a). We see by inspection that $Z_{TH} = R/(sCR + 1)$. Thus, if subnetwork A is removed from Figure 7.30 and replaced by the Thévenin equivalent network, the voltages and currents in subnetwork B remain unchanged.

It is assumed that B in Figure 7.28 has no dependent sources that depend on voltages or currents in A, although dependent sources in B can depend on voltages and currents in B. However, A can have dependent sources, and these dependent sources create a modification in the procedure for finding the Thévenin equivalent network. There may be dependent sources in A that depend on voltages and currents that also exist in A. We call these dependent sources Case I-dependent sources. There may also be dependent sources in A that depend on voltages and currents in B, and we label these sources as Case II-dependent sources. Then, the procedure for finding the Thévenin equivalent network is:

V_{TH} is the voltage across the terminals of Figure 7.29(a). The voltages and currents that Case I-dependent sources depend on must be eliminated from the expression for V_{TH} unless they happen to be the voltages of independent voltage sources or the currents of independent current sources in A. Otherwise, the expression for V_{TH} would not be a finished solution. However, Case II-dependent sources are handled as if they were independent sources. That is, Case II-dependent sources are included in the results for V_{TH} just as independent sources would be.

Z_{TH} is the impedance looking into the terminals in Figure 7.29(b). In this fictitious network, independent sources are properly removed and *Case II-dependent sources are properly removed.* Case I-dependent sources remain in the network and influence the result for the Thévenin impedance. The finished solution for Z_{TH} is a function only of the parameters of the network in Figure 7.29(b) which are Rs, Ls, Cs, Ms (there may be inductive coupling between coils in this network), and the coefficients of the Case I-dependent sources.

Thus, Case II-dependent sources, sources that depend on voltages or currents in subnetwork B, are uniformly treated as if they were independent sources in finding the Thévenin equivalent network. Some examples will clarify the issue.

Example 10. Find the Thévenin equivalent network for subnetwork A in Figure 7.32. Assume the initial current through the inductor is zero.

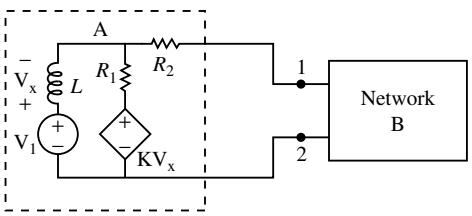


FIGURE 7.32 Network for Example 10.

Solution. There is one independent source and one Case I-dependent source. Figure 7.33(a) depicts the fictitious network to be analyzed to obtain $V_{TH.}$ No current is flowing through R_2 in this figure, therefore, we can write $V_{TH} = V_1 - V_x$. To eliminate V_x from our intermediate expression for V_{TH}, we can use the results of the enhanced voltage divider to write:

$$V_{TH} = \frac{V_1 R_1 + sKL\left(V_1 - V_{TH}\right)}{R_1 + sL}$$

The finished solution for V_{TH} is

$$V_{TH} = V_1 \frac{sKL + R_1}{(K+1)sL + R_1}$$

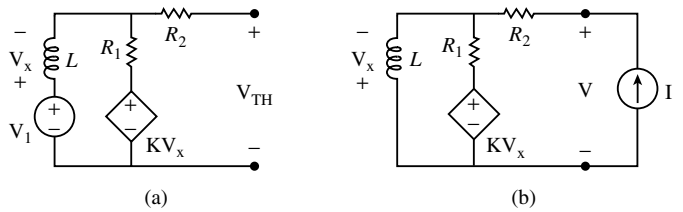


FIGURE 7.33 (a) Abstracted network for finding V_{TH}. (b) Abstracted network for finding Z_{TH}.

Z_{TH} is obtained from Figure 7.33(b) where a current source excitation is shown already applied to the fictitious network. Two node equations, with unknown node voltages V and $-V_x$, enable us to obtain I in terms of V while eliminating V_x. We also note that Z_{TH} consists of resistor R_2 in series with some unknown impedance, so we could remove R_2 (replaced it by a short) if we remember to add it back later. The finished result for Z_{TH} is

$$Z_{TH} = R_2 + \frac{sLR_1}{(K+1)sL + R_1}$$

The following example involves a network having a Case II-dependent source. □

Example 11. Find the Thévenin equivalent network for subnetwork A in the network illustrated in Figure 7.34. In this instance, subnetwork B is outlined explicitly.

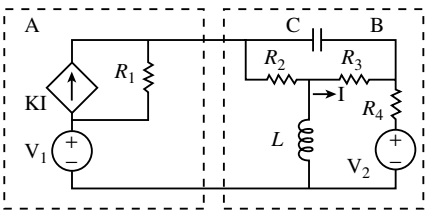


FIGURE 7.34 Network for Example 11.

Solution. Subnetwork A contains one independent source and one Case II-dependent source. Figure 7.35(a) is the abstracted network for finding V_{TH}. Thus,

$$V_{TH} = V_1 + KIR_1$$

Then, both V_1 and the dependent source KI are deleted from Figure 7.35(a) to obtain Figure 7.35(b), the network used for finding Z_{TH}. Thus, $Z_{TH} = R_1$.

Of course, the subnetwork for which the Thévenin equivalent is being determined may have both Case I- and Case II-dependent sources, but these sources can be handled concurrently using the procedures given previously.

Special conditions can arise in the application of Thévenin's theorem. One condition is $Z_{TH} = 0$ and the other is $V_{TH} = 0$. The conditions for which Z_{TH} is zero are:

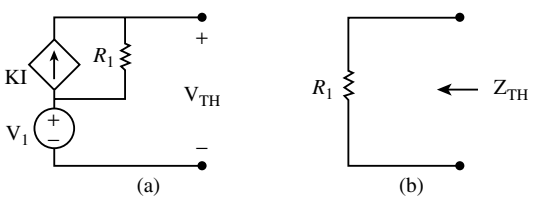


FIGURE 7.35 (a) Network used to find V_{TH}. (b) Network for finding Z_{TH}.

1. If the circuit (subnetwork A) for which the Thévenin equivalent is being determined has an independent voltage source connected between terminals 1 and 2, then $Z_{TH} = 0$. Figure 7.36(a) illustrates this case.
2. If subnetwork A has a *dependent* voltage source connected between terminals 1 and 2, then Z_{TH} is zero *provided* neither of the port variables associated with the port formed by terminals 1 and 2 is coupled back into the network. Figure 7.36(b) is a subnetwork A for which Z_{TH} is zero. However, Figure 7.36(c) depicts a subnetwork A in which the port variable I is coupled back into A by the dependent source $K_1 I$. If I is considered to be a variable of subnetwork A so that $K_1 I$ is a Case I-dependent source, then Z_{TH} is not zero.

The other special condition, $V_{TH} = 0$, occurs if subnetwork A contains only Case I-dependent sources, no independent sources, and no Case II-dependent sources. An example of such a network is given in Figure 7.36(d). With subnetwork B disconnected, subnetwork A is a dead network, and its Thévenin voltage is zero.

The network in Figure 7.36(c) is of interest because the dependent source $K_1 I$ can be considered as a Case I- or a Case II-dependent source hinging on whether I is considered a variable of subnetwork A or B.

Example 12. Solve for I in Figure 7.36(c) using two versions of the Thévenin equivalent for subnetwork A. For the first version, consider I to be associated with A, and therefore both dependent sources are Case I-dependent sources. In the second version, consider I to be associated with B.

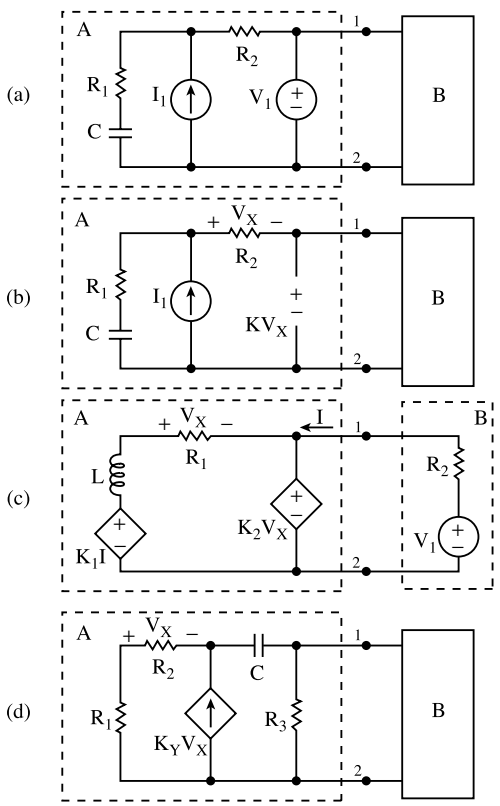


FIGURE 7.36 Special cases of Thévenin's theorem. (a, b) Z_{TH} equals zero. (c) A port variable is coupled back into A. (d) V_{TH} is zero.

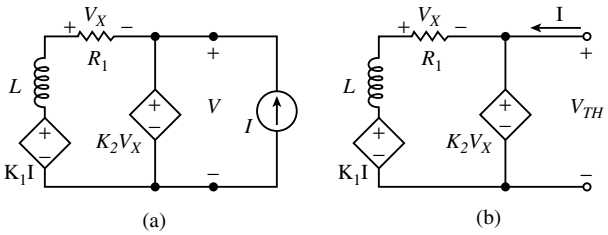


FIGURE 7.37 (a) Network for finding Z_{TH} when both sources are Case I-dependent sources. (b) Network for finding V_{TH} when a Case II-dependent source exists in the network.

Solution. If I is considered as associated with A, then V_{TH} is zero by inspection because A contains only Case I-dependent sources. Figure 7.37(a) depicts subnetwork A with a current excitation applied in order to determine Z_{TH}. Clearly, $V = K_2V_x$. Also, writing a loop equation in the loop encompassed by the two dependent sources, we obtain

$$K_1 I - \frac{V_x}{R_1}\left(sL + R_1\right) = V$$

Eliminating V_x, we have

$$\frac{V}{I} = Z_{TH} = \frac{K_1K_2R_1}{sL + R_1\left(K_2 + 1\right)}$$

Once Z_{TH} is obtained, it is an easy matter to write from Figure 7.36(c):

$$I = \frac{V_1}{R_2 + Z_{TH}} = \frac{V_1\left[sL + R_1\left(K_2 + 1\right)\right]}{sLR_2 + R_1R_2\left(K_2 + 1\right) + K_1K_2R_1}$$

If I is associated with B, then V_{TH} is found from the network in Figure 7.37(b) with the source K_1I treated as if it were independent. The equation for V_{TH} may contain the variable I, but V_x must be eliminated from the finished expression for V_{TH}. We obtain

$$V_{TH} = I\frac{K_1K_2R_1}{sL + R_1\left(K_2 + 1\right)}$$

Also, Z_{TH} is zero because if K_1I is removed, subnetwork A reduces to a network with only a Case I-dependent source and a port variable is not coupled back into the network. Finally, I can be written as:

$$I = \frac{V_1 - V_{TH}}{R_2}$$

which yields the same result for I as was found previously. □

The following example illustrates the interplay that can be achieved among these theorems and source transformations.

Example 13. Find V_0/V_1 for the bridged T network shown in Figure 7.38.

Solution. The application of source transformation eight to the network yields the ladder network in Figure 7.39(a). Thévenin's theorem is particularly useful in analyzing ladder networks. If it is applied to

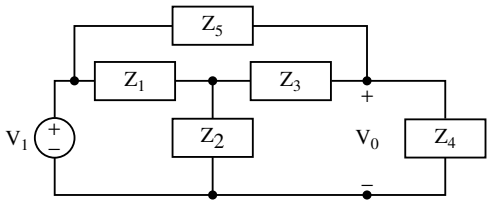


FIGURE 7.38 Network for Example 13.

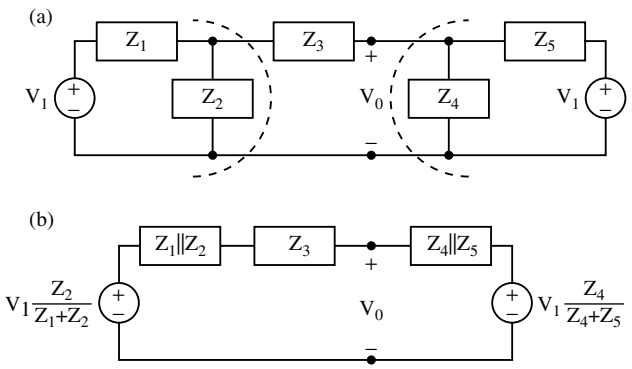


FIGURE 7.39 (a) Results after applying source transformation eight to the network shown in Figure 7.38. (b) Results of two applications of Thévenin's theorem.

the left and right sides of the network, taking care not to obscure the nodes between which V_0 exists, we obtain the single loop network in Figure 7.39(b). Then, using a voltage divider, we obtain

$$\frac{V_0}{V_1} = \frac{Z_4\left[Z_1\left(Z_2+Z_3\right)+Z_2Z_3+Z_2Z_5\right]}{\left(Z_1+Z_2\right)\left(Z_3Z_4+Z_3Z_5+Z_4Z_5\right)+Z_1Z_2\left(Z_4+Z_5\right)}$$

Norton's Theorem

If a source transformation is applied to the Thévenin equivalent network consisting of V_{TH} and Z_{TH} in Figure 7.28(b), then a Norton equivalent network, illustrated in Figure 7.40(a), is obtained. The current source $I_{sc} = V_{TH}/Z_{TH}$, $Z_{TH} \neq 0$. If Z_{TH} equals zero in Figure 7.28(b), then the Norton equivalent network does not exist. The subscripts "*sc*" on the current source stand for short circuit and indicate a procedure for finding the value of this current source. To find I_{sc} for subnetwork A in Figure 7.28(a), we disconnet subnetwork B and place a short circuit between nodes 1 and 2 of subnetwork A. Then, I_{sc} is the current flowing through the short circuit in the direction indicated in Figure 7.40(b). I_{sc} is zero if subnetwork A has only Case I-dependent sources and no other sources. Z_{TH} is found in the same manner as for Thévenin's theorem.

It is sometimes more convenient to find I_{sc} and V_{TH} instead of Z_{TH}.

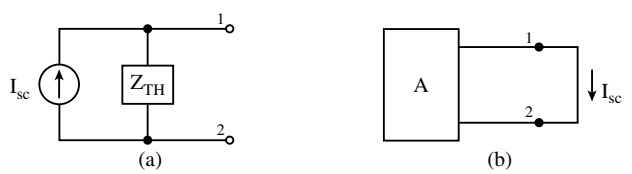


FIGURE 7.40 (a) Norton equivalent network. (b) Reference direction for I_{sc}.

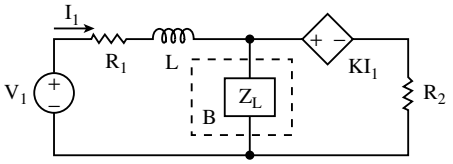


FIGURE 7.41 Network for Example 14.

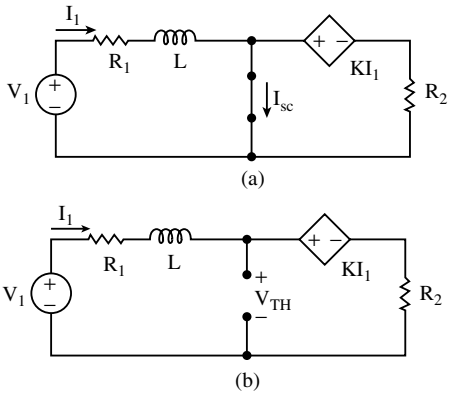


FIGURE 7.42 (a) Network for finding $I_{sc.}$ (b) Network for V_{TH}.

Example 14. Find the Norton equivalent for the network "seen" by Z_L in Figure 7.41. That is, Z_L is subnetwork B and the rest of the network is A, and we wish to find the Norton equivalent network for A.

Solution. Figure 7.42(a) is the network with Z_L replaced by a short circuit. An equation for I_{sc} can be obtained quickly using superposition. This yields

$$I_{sc} = I_1 + \frac{KI_1}{R_2}$$

but I_1 must be eliminated from this equation. I_1 is obtained as: $I_1 = V_1/(sL + R_1)$. Thus,

$$I_{sc} = \frac{V_1\left(1 + \dfrac{K}{R_2}\right)}{sL + R_1}$$

V_{TH} is found from the network shown in Figure 7.42(b). The results are:

$$V_{TH} = \frac{V_1(K + R_2)}{sL + R_1 + R_2 + K}$$

Z_{TH} can be found as V_{TH}/I_{sc}.

Thévenin's and Norton's Theorems and Network Equations

Thévenin's and Norton's theorems can be related to loop and node equations. Here, we examine the relationship to loop equations by means of the LLFT network in Figure 7.43. Assume that the network N in Figure 7.43 has n independent loops with all the loop currents chosen in the same direction. Without loss of generality, assume that only one loop current, say I_1, flows through Z_L as shown so that Z_L appears

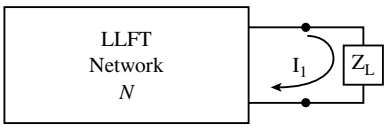


FIGURE 7.43 An LLFT network N with *n* independent loops.

in only one loop equation. For simplicity, assume that no dependent sources or inductive couplings in N exist, and that all current sources have been source-transformed so that only voltage source excitations remain. Then the loop equations are

$$\begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} = \begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1n} \\ z_{21} & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ z_{n1} & z_{n2} & \cdots & z_{nn} \end{bmatrix} \begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} \tag{7.32}$$

where V_i, $i = 1, 2, \ldots, n$, is the sum of all voltage sources in the i*th* loop. Thus, V_i may consist of several terms, some of which may be negative depending on whether a voltage source is a voltage rise or a voltage drop. Also, the impedances z_{ij} are given by

$$z_{ij} = \pm\left[R_{ij} + sL_{ij} + \frac{D_{ij}}{s}\right] \tag{7.33}$$

where $i, j = 1, 2, \ldots, n$, and where the plus sign is taken if $i = j$, and the minus sign is used if $i \neq j$. R_{ij} is the sum of the resistances in loop i if $i = j$, and R_{ij} is the sum of the resistances common to loops i and j if $i \neq j$. Similar statements apply to the inductances L_{ij} and to the reciprocal capacitances D_{ij}. The currents I_i, $i = 1, 2, \ldots, n$, are the unknown loop currents.

Note that Z_L is included only in z_{11} so that z_{11} can be written as $z_{11} = z_{11A} + Z_L$, where z_{11A} is the sum of all the impedances around loop one except Z_L. Solving for I_1 using Cramer's rule, we have

$$I_1 = \frac{\begin{vmatrix} V_1 & z_{12} & \cdots & z_{1n} \\ V_2 & z_{22} & \cdots & z_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ V_n & z_{n2} & \cdots & z_{nn} \end{vmatrix}}{\Delta} \tag{7.34}$$

where Δ is the determinant of the loop impedance matrix. Thus, we can write

$$I_1 = \frac{V_1\Delta_{11} + V_2\Delta_{21} + \cdots + V_n\Delta_{n1}}{z_{11}\Delta_{11} + z_{21}\Delta_{21} + \cdots + z_{n1}\Delta_{n1}} \tag{7.35}$$

or

$$I_1 = \frac{V_1 + V_2\dfrac{\Delta_{21}}{\Delta_{11}} + \cdots + V_n\dfrac{\Delta_{n1}}{\Delta_{11}}}{z_{11} + z_{21}\dfrac{\Delta_{21}}{\Delta_{11}} + \cdots + z_{n1}\dfrac{\Delta_{n1}}{\Delta_{11}}} \tag{7.36}$$

where Δ_{ij} are cofactors of the loop impedance matrix. Then, forming the product of I_1 and Z_L, we have:

$$I_1 Z_L = \frac{Z_L\left(V_1 + V_2\frac{\Delta_{21}}{\Delta_{11}} + \cdots + V_n\frac{\Delta_{n1}}{\Delta_{11}}\right)}{Z_L + z_{11A} + z_{21}\frac{\Delta_{21}}{\Delta_{11}} + \cdots + z_{n1}\frac{\Delta_{n1}}{\Delta_{11}}} \tag{7.37}$$

If we take the limit of I_1Z_L as Z_L approaches infinity, we obtain the "open circuit" voltage V_{TH}. That is,

$$\lim_{Z_L\to\infty} I_1 Z_L = V_{TH} = \left(V_1 + V_2\frac{\Delta_{21}}{\Delta_{11}} + \cdots + V_n\frac{\Delta_{n1}}{\Delta_{11}}\right) \tag{7.38}$$

and if we take the limit of I_1 as Z_L approaches zero, we obtain the "short circuit" current I_{sc}:

$$\lim_{Z_L\to 0} I_1 = I_{sc} = \frac{V_{TH}}{z_{11A} + z_{21}\frac{\Delta_{21}}{\Delta_{11}} + \cdots + z_{n1}\frac{\Delta_{n1}}{\Delta_{11}}} \tag{7.39}$$

Finally, the quotient of V_{TH} and I_{sc} yields:

$$\frac{V_{TH}}{I_{sc}} = Z_{TH} = z_{11A} + z_{21}\frac{\Delta_{21}}{\Delta_{11}} + \cdots + z_{n1}\frac{\Delta_{n1}}{\Delta_{11}} \tag{7.40}$$

If network N contains coupled inductors (but not coupled to Z_L), then some elements of the loop impedance matrix may be modified in value and sign. If network N contains dependent sources, then auxiliary equations can be written to express the quantities on which the dependent sources depend in terms of the independent excitations and/or the unknown loop currents. Thus, dependent sources may modify the elements of the loop impedance matrix in value and sign, and they may modify the elements of the excitation column matrix $[V_i]$. Nevertheless, we can obtain expressions similar to those obtained previously for V_{TH} and I_{sc}. Of course, we must exclude from this illustration dependent sources that depend on the voltage across Z_L because they violate the assumption that Z_L appears in only one loop equation and are beyond the scope of this discussion.

The π-T Conversion

The π-T conversion is employed for the simplification of circuits, especially in power systems analysis. The "π" refers to a circuit having the topology shown in Figure 7.44. In this figure, the left-most and right-most loop currents have been chosen to coincide with the port currents for convenience of notation only.

A circuit having the topology shown in Figure 7.45 is referred to as a "T" or as a "Y." We wish to determine equations for Z_1, Z_2, and Z_3 in terms of Z_A, Z_B, and Z_C so that the π can be replaced by a T

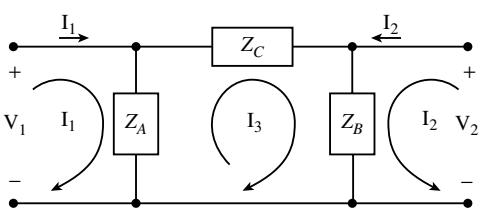


FIGURE 7.44 A π network shown with loop currents.

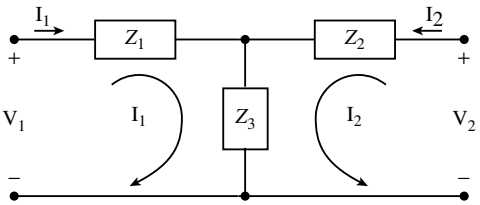


FIGURE 7.45 A T network.

without affecting any of the port variables. In other words, if an overall circuit contains a π subcircuit, we wish to replace the π subscript with a T subscript without disturbing any of the other voltages and currents within the overall circuit. To determine what Z_1, Z_2, and Z_3 should be, we first write loop equations for the π network. The results are:

$$V_1 = I_1 Z_A - I_3 Z_A \tag{7.41}$$

$$V_2 = I_2 Z_B + I_3 Z_B \tag{7.42}$$

$$0 = I_3\left(Z_A + Z_B + Z_C\right) - I_1 Z_A + I_2 Z_B \tag{7.43}$$

But the T circuit has only two loop equations given by:

$$V_1 = I_1\left(Z_1 + Z_3\right) + I_2 Z_3 \tag{7.44}$$

$$V_2 = I_1 Z_3 + I_2\left(Z_2 + Z_3\right) \tag{7.45}$$

We must eliminate one of the loop equations for the π circuit, and so we solve for I_3 in (7.43) and substitute the result into (7.41) and (7.42) to obtain:

$$V_1 = I_1\left[\frac{Z_A\left(Z_B + Z_C\right)}{Z_A + Z_B + Z_C}\right] + I_2\left[\frac{Z_A Z_B}{Z_A + Z_B + Z_C}\right] \tag{7.46}$$

$$V_2 = I_1\left[\frac{Z_A Z_B}{Z_A + Z_B + Z_C}\right] + I_2\left[\frac{Z_B\left(Z_A + Z_C\right)}{Z_A + Z_B + Z_C}\right] \tag{7.47}$$

From a comparison of the coefficients of the currents in (7.46) and (7.47) with those in (7.44) and (7.45), we obtain the following relationships.

Replacing π with T

$$Z_1 = \frac{Z_A Z_C}{S_Z};\quad Z_2 = \frac{Z_B Z_C}{S_Z};\quad Z_3 = \frac{Z_A Z_B}{S_Z} \tag{7.48}$$

where

$$S_Z = Z_A + Z_B + Z_C$$

We can also replace a T network by a π network. To do this we need equations for Z_A, Z_B, and Z_C in terms of Z_1, Z_2, and Z_3. The required equations can be obtained algebraically from (7.48).

From T to π

$$Z_A = Z_1 + Z_3 + \frac{Z_1 Z_3}{Z_2}; \quad Z_B = Z_2 + Z_3 + \frac{Z_2 Z_3}{Z_1}; \quad Z_C = Z_1 + Z_2 + \frac{Z_1 Z_2}{Z_3} \tag{7.49}$$

Reciprocity

If an LLFT network contains only Rs, Ls, Cs, and transformers but contains no dependent sources, then its loop impedance matrix is symmetrical with respect to the main diagonal. That is, if z_{ij} is an element of the loop impedance matrix (see (7.17)), occupying the position at row i and column j, then $z_{ji} = z_{ij}$, where z_{ji} occupies the position at row j and column i. Such a network has the property of reciprocity and is termed a reciprocal network.

Assume that a reciprocal network, depicted in Figure 7.46, has m loops and is in the zero state. It has only one excitation — a voltage source in loop j. To solve for the loop current in loop k, we write the loop equations:

$$\begin{bmatrix} z_{11} & z_{12} & \cdots & z_{1m} \\ z_{21} & z_{22} & \cdots & z_{2m} \\ \vdots & \vdots & \ddots & \vdots \\ z_{j1} & z_{j2} & \cdots & z_{jm} \\ z_{k1} & z_{k2} & \cdots & z_{km} \\ \vdots & \vdots & \ddots & \vdots \\ z_{m1} & z_{m2} & \cdots & z_{mm} \end{bmatrix} \begin{bmatrix} I_2 \\ I_2 \\ \vdots \\ I_j \\ I_k \\ \vdots \\ I_m \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ V_j \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{7.50}$$

The column excitation matrix has only one nonzero entry. To determine I_k using Cramer's rule, we replace column k by the excitation column and then expand along this column. Only one nonzero term is in the column, therefore, we obtain a single term for I_k:

$$I_k = V_j \frac{\Delta_{jk}}{\Delta} \tag{7.51}$$

where Δ_{jk} is the cofactor, and Δ is the determinant of the loop impedance matrix.

Next, we replace the voltage source by a short circuit in loop j, cut the wire in loop k, and insert a voltage source V_k. Figure 7.47 outlines the modifications to the circuit. Then, we solve for I_j obtaining

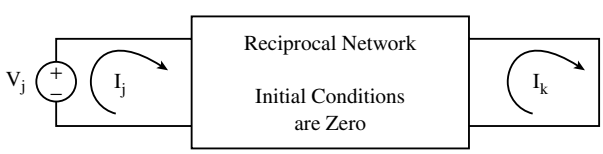


FIGURE 7.46 A reciprocal network with m independent loops.

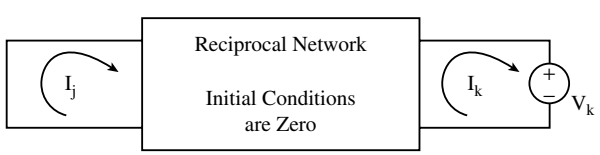


FIGURE 7.47 Interchange of the ports of excitation in the network in Figure 7.46.

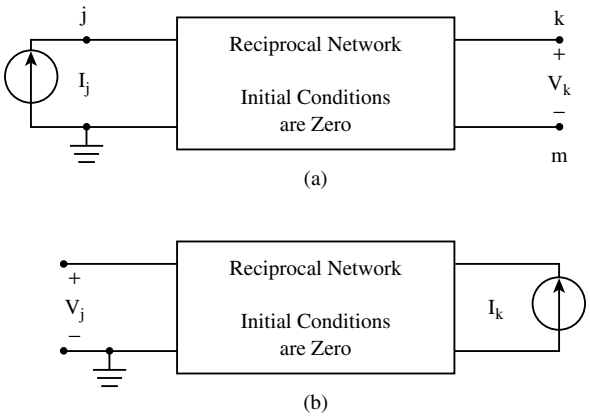


FIGURE 7.48 (a) Reciprocal ungrounded network with a current source excitation. (b) Interchange of the ports of excitation and response.

$$I_j = V_k \frac{\Delta_{kj}}{\Delta} \tag{7.52}$$

Because the network is reciprocal, $\Delta_{jk} = \Delta_{kj}$ so that

$$\frac{I_k}{V_j} = \frac{I_j}{V_k} \tag{7.53}$$

Eq. (7.53) is the statement of reciprocity for the network in Figures 7.46 and 7.47 with the excitations shown.

Figure 7.48(a) is a reciprocal network with a current excitation applied to node *j* and a voltage response, labeled V_k, taken between nodes *k* and *m*. We assume the network has *n* independent nodes plus the ground node indicated and is not a grounded network (does not have a common connection between the input and output ports shown). If we write node equations to solve for V_k in Figure 7.48(a) and use Cramer's rule, we have:

$$V_k = I_j \frac{\Delta'_{jk} - \Delta'_{jm}}{\Delta'} \tag{7.54}$$

where the primes indicate node-basis determinants. Then, we interchange the ports of excitation and response as depicted in Figure 7.48(b). If we solve for V_j in Figure 7.48(b), we obtain

$$V_j = I_k \frac{\Delta'_{kj} - \Delta'_{mj}}{\Delta'} \tag{7.55}$$

Because the corresponding determinants in (7.54) and (7.55) are equal because of reciprocity, we have:

$$\frac{V_k}{I_j} = \frac{V_j}{I_k} \tag{7.56}$$

Note that the excitations and responses are of the opposite type in Figures 7.46 and 7.48. The results obtained in (7.53) and (7.56) do not apply if the excitation and response are both voltages or both currents because when the ports of excitation and response are interchanged, the impedance levels of the network are changed [2].

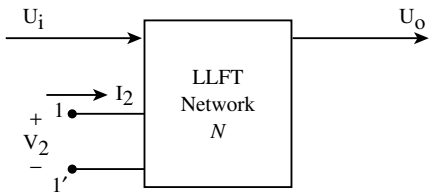


FIGURE 7.49 An arbitrary LLFT network.

Middlebrook's Extra Element Theorem

Middlebrook's extra element theorem is useful in developing tests for analog circuits and for predicting the effects that parasitic elements may have on a circuit. This theorem has two versions: the parallel version and the series version. Both versions present the results of connecting an extra network element in the circuit as the product of the network function obtained without the extra element times a correction factor. This is a particularly convenient form for the results because it shows exactly the effects of the extra element on the network function.

Parallel Version. Consider an arbitrary LLFT network having a transfer function $A_1(s)$. In the parallel version of the theorem, an impedance is added between any two independent nodes of the network. The modified transfer function is then obtained as $A_1(s)$ multiplied by a correction factor. Figure 7.49 is an arbitrary network in which the quantities U_i and U_0 represent a general input and a general output, respectively, whether they are voltages or currents. The extra element is to be connected between terminals 1 and 1′ in Figure 7.49, and the port variables for this port are V_2 and I_2.

We can write:

$$\begin{aligned} U_o &= A_1 U_i + A_2 I_2 \\ V_2 &= B_1 U_i + B_2 I_2 \end{aligned} \tag{7.57}$$

where

$$\begin{aligned} A_1 &= \left.\frac{U_o}{U_i}\right|_{I_2=0} \quad A_2 = \left.\frac{U_o}{I_2}\right|_{U_i=0} \\ B_1 &= \left.\frac{V_2}{U_i}\right|_{I_2=0} \quad B_2 = \left.\frac{V_2}{I_2}\right|_{U_i=0} \end{aligned} \tag{7.58}$$

Note that A_1 is assumed to be known.

The extra element Z to be added across terminals 1 and 1′ is depicted in Figure 7.50. It can be described as $Z = V_2/(-I_2)$ which yields $I_2 = V_2/(-Z)$. Substituting this expression for I_2 into (7.57) results in:

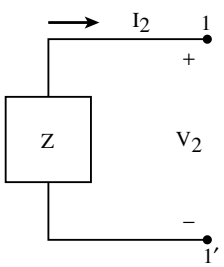


FIGURE 7.50 The extra element Z.

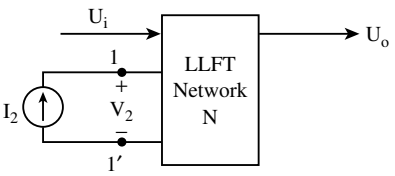


FIGURE 7.51 Network of Figure 7.49 with two excitations applied.

$$U_o = A_1 U_i + A_2\left(\frac{-V_2}{Z}\right)$$
$$V_2\left(1+\frac{B_2}{Z}\right) = B_1 U_i \tag{7.59}$$

After eliminating V_2 and solving for U_o/U_i, we obtain:

$$\frac{U_o}{U_i} = A_1\left[\frac{1+\frac{1}{Z}\left(\frac{A_1B_2 - A_2B_1}{A_1}\right)}{1+\frac{B_2}{Z}}\right] \tag{7.60}$$

Next, we provide physical interpretations for the terms in (7.60). Clearly, B_2 is the impedance seen looking into the network between terminals 1 and 1′ with $U_i = 0$. Thus, rename $B_2 = Z_d$ where d stands for "dead network."

To find a physical interpretation of $(A_1B_2 - A_2B_1)/A_1$, examine the network in Figure 7.51. Here, two excitations are applied to the network, namely U_i and I_2. Simultaneously adjust both inputs so as to null output U_o. Thus, with $U_o = 0$, we have from (7.57),

$$U_i = \frac{-A_2}{A_1} I_2 \tag{7.61}$$

Substituting this result into the equation for V_2 in (7.57), we have:

$$V_2 = B_1\left(\frac{-A_2}{A_1}\right) I_2 + B_2 I_2 \tag{7.62}$$

or

$$\left.\frac{V_2}{I_2}\right|_{U_o=0} = \frac{A_1B_2 - A_2B_1}{A_1}$$

Because the quantity $(A_1B_2 - A_2B_1)/A_1$ is the ratio of V_2 to I_2 with the output *nulled*, we rename this quantity as Z_N. Then rewriting (7.60) with Z_d and Z_N, we have:

$$\frac{U_o}{U_i} = A_1\left[\frac{1+\frac{Z_N}{Z}}{1+\frac{Z_d}{Z}}\right] \tag{7.63}$$

Equation (7.63) demonstrates that the results of connecting the extra element Z into the circuit can be expressed as the product of A_1, which is the network function with Z set to infinity, times a correction factor given in the brackets in (7.63).

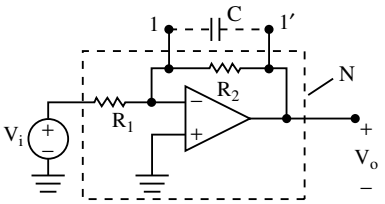


FIGURE 7.52 Ideal op amp circuit for Example 15.

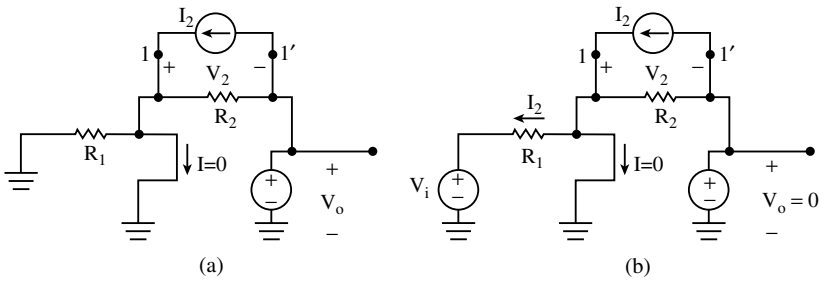


FIGURE 7.53 (a) Network for finding Z_d. (b) Network used to determine Z_N.

Example 15. Use the parallel version of Middlebrook's extra element theorem to find the voltage transfer function of the ideal op amp circuit in Figure 7.52 when a capacitor C is connected between terminals 1 and 1′.

Solution. With the capacitor not connected, the voltage transfer function is

$$\left.\frac{V_0}{V_i}\right|_{Z=\infty} = -\frac{R_2}{R_1} = A_1$$

Next, we determine Z_d from the circuit illustrated in Figure 7.53(a), where a model has been included for the ideal op amp, the excitation V_i has been properly removed, and a current excitation I_2 has been applied to the port formed by terminals 1 and 1′. Because no voltage flows across R_1 in Figure 7.53(a), no current flows through it, and all the current I_2 flows through R_2. Thus, $V_2 = I_2R_2$, and $Z_d = R_2$. We next find Z_N from Figure 7.53(b). We observe in this figure that the right end of R_2 is zero volts above ground because V_i and I_2 have been adjusted so that V_0 is zero. Furthermore, the left end of R_2 is zero volts above ground because of the virtual ground of the op amp. Thus, zero is current flowing through R_2, and so V_2 is zero. Consequently, $Z_N = V_2/I_2 = 0$. Following the format of (7.63), we have:

$$\frac{V_0}{V_i} = -\frac{R_2}{R_1}\left(\frac{1}{1+sCR_2}\right)$$

Note that for V_0 to be zero in Figure 7.53(b), V_i and I_2 must be adjusted so that $V_i/R_1 = -I_2$, although this information was not needed to work the example. □

Series Version. The series version of the theorem allows us to cut a loop of the network, add an impedance Z in series, and obtain the modified network function as $A_1(s)$ multiplied by a correction factor. A_1 is the network function when $Z = 0$. Figure 7.54 is an LLFT network with part of a loop illustrated explicitly. The quantities U_i and U_0 represent a general input and a general output, respectively, whether they be a voltage or a current. Define

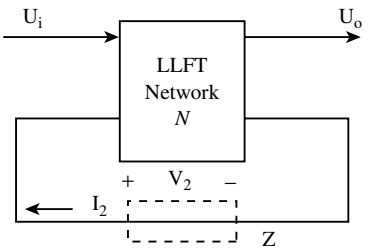


FIGURE 7.54 LLFT network used for the series version of Middlebrook's extra element theorem.

$$A_1 = \left.\frac{U_0}{U_i}\right|_{V_2=0} \quad A_2 = \left.\frac{U_0}{V_2}\right|_{U_i=0}$$
$$B_1 = \left.\frac{I_2}{U_i}\right|_{V_2=0} \quad B_2 = \left.\frac{I_2}{V_2}\right|_{U_i=0} \tag{7.64}$$

where V_2 and I_2 are depicted in Figure 7.54, and A_1 is assumed to be known. Then using superposition, we have:

$$U_0 = A_1 U_i + A_2 V_2$$
$$I_2 = B_1 U_i + B_2 V_2 \tag{7.65}$$

The impedance of the extra element Z can be described by $Z = V_2/(-I_2)$ so that $V_2 = -I_2 Z$. Substituting this relation for V_2 into (7.65) and eliminating I_2, we have:

$$\frac{U_0}{U_i} = A_1 \left[\frac{1 + Z\left(B_2 - B_1 \dfrac{A_2}{A_1}\right)}{1 + B_2 Z}\right] \tag{7.66}$$

Again, as we did for the parallel version of the theorem, we look for physical interpretations of the quantities in the square bracket in (7.66). From (7.65) we see that B_2 is the admittance looking into the port formed by cutting the loop in Figure 7.54 with $U_i = 0$. This is depicted in Figure 7.55(a). Thus, B_2 is the admittance looking into a dead network, and so let $B_2 = 1/Z_d$.

To find a physical interpretation of the quantity $(A_1B_2 - A_2B_1)/A_1$, we examine Figure 7.55(b) in which both inputs, V_2 and U_i, are adjusted to null the output U_0. From (7.65) with $U_o = 0$, we have:

$$U_i = -\frac{A_2}{A_1} V_2 \tag{7.67}$$

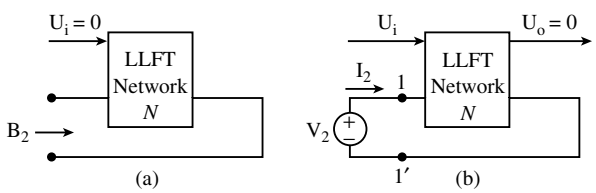


FIGURE 7.55 (a) Looking into the network with U_i equal zero. (b) U_i and V_2 are simultaneously adjusted to null the output U_0.

Then, eliminating U_i in (7.65) we obtain:

$$\frac{A_1B_2 - A_2B_1}{A_1} = \left.\frac{I_2}{V_2}\right|_{U_0=0} \tag{7.68}$$

Because this quantity is the admittance looking into the port formed by terminals 1 and 1′ in Figure 7.55(b) with U_o nulled, rename it as $1/Z_n$. Thus, from (7.66) we can write

$$\frac{U_o}{U_i} = A_1\left[\frac{1+\dfrac{Z}{Z_N}}{1+\dfrac{Z}{Z_d}}\right] \tag{7.69}$$

Eq. (7.69) is particularly convenient for determining the effects of adding an impedance Z into a loop of a network.

Example 16. Use the series version of Middlebrook's extra element theorem to determine the effects of inserting a capacitor C in the location indicated in Figure 7.56.

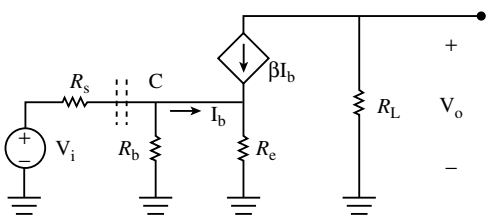


FIGURE 7.56 Network for Example 16.

Solution. The voltage transfer function for the network without the capacitor is found to be:

$$A_1 = \left.\frac{V_0}{V_i}\right|_{Z=0} = \frac{-\beta R_L}{R_s + (\beta+1)R_e\left(1+\dfrac{R_s}{R_b}\right)}$$

Next, we find Z_d from Figure 7.57(a). This yields:

$$Z_d = \frac{V}{I} = R_s + \left[R_b \| (\beta+1)R_e\right]$$

The impedance Z_n is found from Figure 7.57(b) where the two input sources, V_i and V, are adjusted so that V_0 equals zero. If V_0 equals zero, then βI_b equals zero because no current flows through R_L. Thus, I_b equals zero, which implies that V_{Re}, the voltage across R_e as indicated in Figure 7.57(b), is also zero. We see that the null is propagating through the circuit. Continuing to analyze Figure 7.57(b), we see that I_{Rb} is zero so that we conclude that I is zero. Because $Z_N = V/I$, we conclude that Z_N is infinite. Using the format given by (7.69) with $Z = 1/(sC)$, we obtain the result as:

$$\frac{V_0}{V_i} = A_1\left\{\frac{1}{1+\dfrac{1/(sC)}{R_s + \left[R_b \| (\beta+1)R_e\right]}}\right\}$$

□

It is interesting to note that to null the output so that Z_N could be found in Example 16, V_i is set to V, although this fact is not needed in the analysis.

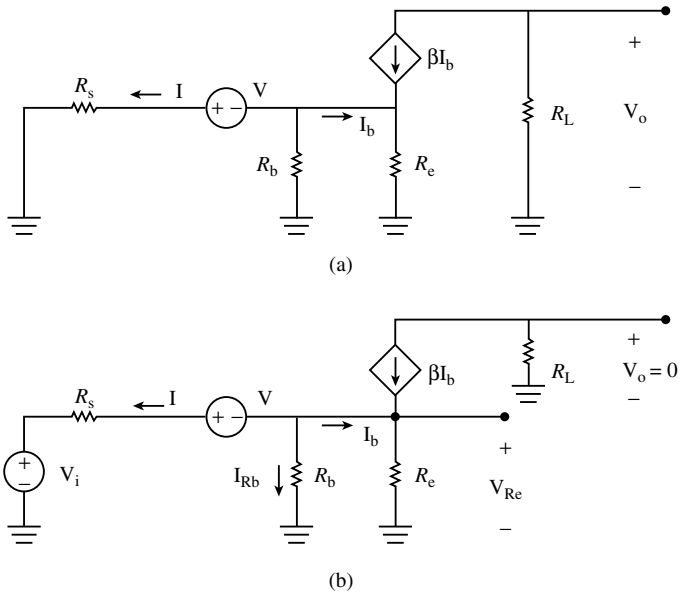


FIGURE 7.57 (a) Network used to obtain Z_d. (b) Network that yields Z_N.

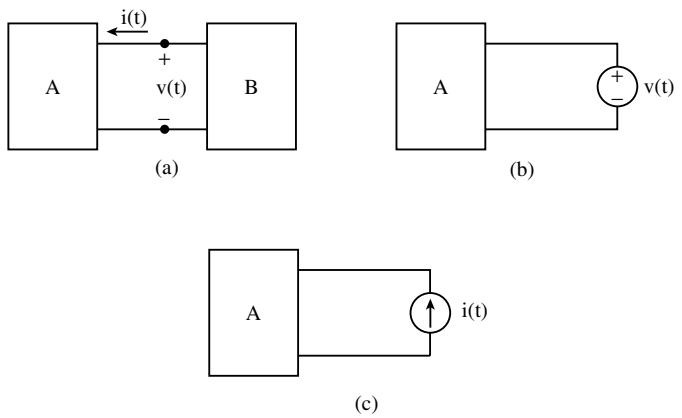


FIGURE 7.58 (a) An LLFT network consisting of two subnetworks A and B connected by two wires. (b) A voltage source can be substituted for subnetwork B if $v(t)$ is known in (a). (c) A current source can be substituted for B if i is a known current.

Substitution Theorem

Figure 7.58(a) is an LLFT network consisting of two subnetworks A and B, which are connected to each other by two wires. If the voltage $v(t)$ is known, the voltages and currents in subnetwork A remain unchanged if a voltage source of value $v(t)$ is substituted for subnetwork B as illustrated in Figure 7.58(b).

Example 17. Determine $i_1(t)$ in the circuit in Figure 7.59. The voltage $v_1(t)$ is known from a previous analysis.

Solution. Because $v_1(t)$ is known, the substitution theorem can be applied to obtain the circuit in Figure 7.60. Analysis of this simplified circuit yields:

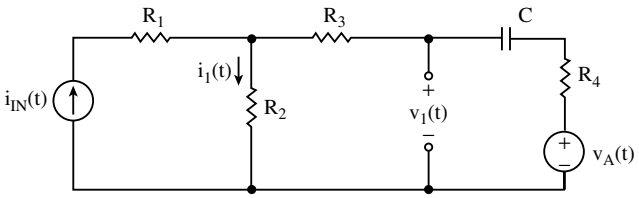


FIGURE 7.59 Circuit for Example 17.

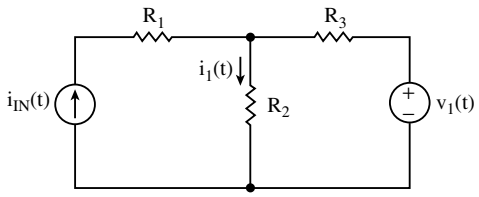


FIGURE 7.60 Circuit that results when the Substitution Theorem is applied to the circuit in Figure 7.59.

$$i_1 = i_{in}\frac{R_3}{R_2 + R_3} + v_1\frac{1}{R_2 + R_3}$$

If the current $i(t)$ is known in Figure 7.58(a), then the substitution in Figure 7.58(c) can be employed.

7.3 Sinusoidal Steady-State Analysis and Phasor Transforms

Sinusoidal Steady-State Analysis

In this section, we develop techniques for analyzing lumped, linear, finite, time invariant (LLFT) networks in the sinusoidal steady state. These techniques are important for analyzing and designing networks ranging from AC power generation systems to electronic filters.

To put the development of sinusoidal steady state analysis in its context, we list the following definitions of responses of circuits:

A. The **zero-input response** is the response of a circuit to its initial conditions when the input excitations are set to zero.
B. The **zero-state response** is the response of a circuit to a given input excitation or set of input excitations when the initial conditions are all set to zero.

The sum of the zero-input response and the zero-state response yields the total response of the system being analyzed. However, the total response can also be decomposed into the **forced response** and the **natural response** if the input excitations are DC, real exponentials, sinusoids, and/or sinusoids multiplied by real exponentials and if the exponent(s) in the input excitation differs from the exponents appearing in the zero-input response. These excitations are very common in engineering applications, and the decomposition of the response into forced and natural components corresponds to the particular and complementary (homogeneous) solutions, respectively, of the linear, constant coefficient, ordinary differential equations that characterize LLFT networks in the time domain. Therefore, we define:

C. The **forced response** is the portion of the total response that has the same exponents as the input excitations.
D. The **natural response** is the portion of the total response that has the same exponents as the zero-input response.

The sum of the forced and natural responses is the total response of the system.

For a strictly stable LLFT network, meaning that the poles of the system transfer function $T(s)$ are confined to the open left-half s-plane (LHP), the natural response must decay to zero eventually. The forced response may or may not decay to zero depending on the excitation and the network to which it is applied, and so it is convenient to define the terms **steady-state response** and **transient response**:

E. The **transient response** is the portion of the response that dies away or decays to zero with time.
F. The **steady-state response** is the portion of the response that does not decay with time.

The sum of the transient and steady state responses is the total response, but a specific circuit with a specific excitation may not have a transient response or it may not have a steady state response. The following example illustrates aspects of these six definitions.

Example 18. Find the total response of the network shown in Figure 7.61, and identify the zero-state, zero-input, forced, natural, transient, and steady-state portions of the response.

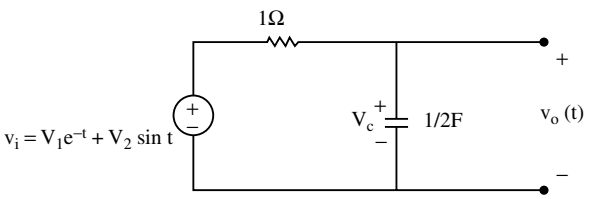


FIGURE 7.61 Circuit for Example 18.

Solution. Note that a nonzero initial condition is represented by V_c. Using superposition and the simple voltage divider concept, we can write:

$$V_0(s) = \left[\frac{V_1}{s+1} + \frac{V_2}{s^2+1}\right]\frac{2}{s+2} + \frac{V_c}{s}\left(\frac{s}{s+2}\right)$$

The partial fraction expansion for $V_0(s)$ is:

$$V_0(s) = \frac{A}{s+1} + \frac{B}{s+2} + \frac{Cs+D}{s^2+1}$$

where

$$A = 2V_1$$

$$B = -2V_1 + 0.4V_2 + V_c$$

$$C = -0.4V_2$$

$$D = 0.8V_2$$

Thus, for $t \geq 0$, $v_0(t)$ can be written as:

$$v_0(t) = 2V_1e^{-t} - 2V_1e^{-2t} + 0.4V_2e^{-2t} + V_ce^{-2t}$$
$$-0.4V_2\cos t + 0.8V_2\sin t$$

With the aid of the angle sum and difference formula

$$\sin(\alpha \pm \beta) = \sin\alpha\cos\beta \pm \cos\alpha\sin\beta$$

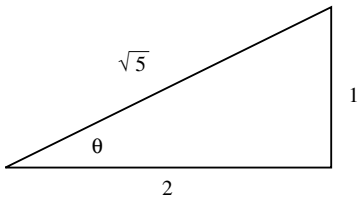


FIGURE 7.62 Sketch for determining the phase angle θ.

and the sketch in Figure 7.62, we can combine the last two terms in the expression for v_0 (t) to obtain:

$$v_0(t) = 2V_1e^{-t} - 2V_1e^{-2t} + 0.4V_2e^{-2t} + V_ce^{-2t}$$
$$+ 0.4\sqrt{5}V_2\sin(t+\theta)$$

where

$$\theta = -\tan^{-1}\left(\frac{1}{2}\right)$$

The terms of $v_0(t)$ are characterized by our definitions as follows:

$$\text{zero-input response} = V_ce^{-2t}$$
$$\text{zero-state response} = 2V_1e^{-t} - 2V_1e^{-2t} + 0.4V_2e^{-2t} + 0.4\sqrt{5}V_2\sin(t+\theta)$$
$$\text{natural response} = \left[-2V_1 + 0.4V_2 + V_c\right]e^{-2t}$$
$$\text{forced response} = 2V_1e^{-t} + 0.4\sqrt{5}V_2\sin(t+\theta)$$
$$\text{transient response} = 2V_1e^{-t} + \left[-2V_1 + 0.4V_2 + V_c\right]e^{-2t}$$
$$\text{steady-state response} = 0.4\sqrt{5}V_2\sin(t+\theta)$$

As can be observed by comparing the preceding terms above with the total response, part of the forced response is the steady state response, and the rest of the forced response is included in the transient response in this example. □

If the generator voltage in the previous example had been $v_i = V_1 + V_2$ sin(t), then there would have been two terms in the steady state response — a DC term and a sinusoidal term. On the other hand, if the transfer function from input to output had a pole at the origin and the excitation were purely sinusoidal, there would also have been a DC term and a sinusoidal term in the steady state response. The DC term would have arisen from the pole at the origin in the transfer function, and therefore would also be classed as a term in the natural response.

Oftentimes, it is desirable to obtain only the sinusoidal steady state response, without having to solve for other portions of the total response. The ability to solve for just the sinusoidal steady state response is the goal of sinusoidal steady state analysis.

The sinusoidal steady state response can be obtained based on analysis of the network using Laplace transforms. Figure 7.63 illustrates an LLFT network that is excited by the voltage sine wave $v_i(t) = V$ sin(ωt), where V is the peak amplitude of the sine wave and ω is the frequency of the sine wave in radians/second. Assume that the poles of the network transfer function $V_0(s)/V_i(s) = T(s)$ are confined to the open left-half s-plane (LHP) except possibly for a single pole at the origin. Then, the forced response of the network is

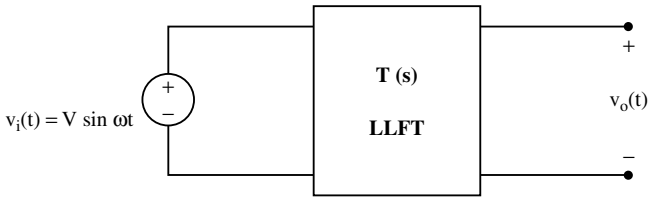


FIGURE 7.63 LLFT network with transfer function T (s).

$$v_{oss}(t) = V\left|T(j\omega)\right|\sin\left(\omega t + \theta\right) \tag{7.70}$$

where the extra subscripts "*ss*" on $v_0(t)$ indicate sinusoidal steady state, and where

$$\theta = \tan^{-1}\left(\frac{\Im T(j\omega)}{\Re T(j\omega)}\right) \tag{7.71}$$

The symbols $\Im$ and $\Re$ are read as "imaginary part of" and "real part of," respectively. In other words, the LLFT network modifies the sinusoidal input signal in only two ways at steady state. The network multiplies the amplitude of the signal by $|T(j\omega)|$ and shifts the phase by θ. If the transfer function of the network is known beforehand, then the sinusoidal steady state portion of the total response can be easily obtained by means of Eqs. (7.70) and (7.71).

To prove (7.70) and (7.71), we assume that $T(s) = V_0(s)/V_i(s)$ in Figure 7.63 is real for s real and that the poles of $T(s)$ are confined to the open LHP except possibly for a single pole at the origin. Without loss of generality, assume the order of the numerator of $T(s)$ is at most one greater than the order of the denominator. Then the transform of the output voltage is

$$V_0(s) = V_i(s)T(s) = \frac{V\omega}{s^2 + \omega^2}T(s)$$

If $V_0(s)$ is expanded into partial fractions, we have:

$$V_0(s) = \frac{A}{s - j\omega} + \frac{B}{s + j\omega} + \text{other terms due to the poles of } T(s)$$

The residue A is

$$A = \left[(s - j\omega)V_o(s)\right]_{s=j\omega} = \left[\frac{V\omega}{s + j\omega}T(s)\right]_{s=j\omega}$$

$$= \frac{V}{2j}T(j\omega)$$

But

$$T(j\omega) = \left|T(j\omega)\right|e^{j\theta} \text{ where } \theta = \tan^{-1}\frac{\Im T(j\omega)}{\Re T(j\omega)}$$

Thus, we can write the residue A as

$$A = \frac{V}{2j}\left|T(j\omega)\right|e^{j\theta}$$

Also, $B = A^*$ where "*" denotes "conjugate," and so

$$B = -\frac{V}{2j}T(-j\omega) = -\frac{V}{2j}T\big((j\omega)^*\big)$$

$$= -\frac{V}{2j}T^*(j\omega) = -\frac{V}{2j}\left|T(j\omega)\right|e^{-j\theta}$$

In the equation for the residue B, we can write $T((j\omega)^*) = T^*(j\omega)$ because of the assumption that $T(s)$ is real for s real (see Property 1 in Section 7.1 on "Properties of LLFT Network Functions").

All other terms in the partial fraction of $V_0(s)$ will yield, when inverse transformed, functions of time that decay to zero except for a term arising from a pole at the origin of $T(s)$. A pole at the origin yields, when its partial fraction is inverse transformed, a DC term that is part of the steady-state solution in the time domain. However, only the first two terms in $V_0(s)$ will ultimately yield a sinusoidal function. We can rewrite these two terms as:

$$V_{oss}(s) = \frac{V}{2j}\left|T(j\omega)\right|\left[\frac{e^{j\theta}}{s-j\omega} - \frac{e^{-j\theta}}{s+j\omega}\right]$$

The extra subscripts "*ss*" denote sinusoidal steady state. The time-domain equation for the sinusoidal steady state output voltage is

$$v_{oss}(t) = \frac{V}{2j}\left|T(j\omega)\right|\left[e^{j\theta}e^{j\omega t} - e^{-j\theta}e^{-j\omega t}\right]$$

$$= V\left|T(j\omega)\right|\sin(\omega t + \theta)$$

where θ is given by (7.71). This completes the proof.

Example 19. Verify the expression for the sinusoidal steady-state response found in the previous example.

Solution. The transfer function for the network in Figure 7.61 is $T(s) = 2/(s + 2)$, and the frequency of the sinusoidal portion of $v_i(t)$ is $\omega = 1$ rad/s. Thus,

$$T(j\omega) = \frac{2}{j+2} = \frac{2}{\sqrt{4+1}}e^{j\theta}$$

where

$$\theta = \tan^{-1}\left(\frac{-2}{4}\right) = -\tan^{-1}\left(\frac{1}{2}\right)$$

If the excitation in Figure 7.63 were $v_i(t) = V \sin(\omega t + \Phi)$, then the sinusoidal steady-state response of the network would be:

$$v_{oss}(t) = V\left|T(j\omega)\right|\sin(\omega t + \Phi + \theta) \tag{7.72}$$

where θ is given by (7.71). Similarly, if the excitation were $v_i(t) = V[\cos(\omega t + \Phi)]$, then the sinusoidal steady-state response would be expressed as:

$$v_{oss}(t) = V\left|T(j\omega)\right|\cos(\omega t + \Phi + \theta) \tag{7.73}$$

with θ again given by (7.71).

Phasor Transforms

In the sinusoidal steady-state analysis of stable LLFT networks, we find that both the inputs and outputs are sine waves of the same frequency. The network only modifies the amplitudes and the phases of the sinusoidal input signals; it does not change their nature. Thus, we need only keep track of the amplitudes and phases, and we do this by using phasor transforms. Phasor transforms are closely linked to Euler's identity:

$$e^{\pm j\omega t} = \cos(\omega t) \pm j\sin(\omega t) \tag{7.74}$$

If, for example, $v_i(t) = V\sin(\omega t + \Phi)$, then we can write $v_i(t)$ as

$$v_i(t) = \mathcal{I}\left[Ve^{j(\omega t+\Phi)}\right] = \mathcal{I}\left[Ve^{j\Phi}e^{j\omega t}\right] \tag{7.75}$$

Similarly, if $v_i(t) = V\cos(\omega t + \Phi)$, then we can write

$$v_i(t) = \mathcal{R}\left[Ve^{j(\omega t+\Phi)}\right] = \mathcal{R}\left[Ve^{j\Phi}e^{j\omega t}\right] \tag{7.76}$$

If we confine our analysis to single-frequency sine waves, then we can drop the imaginary sign and the term $e^{j\omega t}$ in (7.75) to obtain the phasor transform. That is,

$$\wp\left[v_i(t)\right] = \wp\left[V\sin(\omega t + \Phi)\right] = \wp\left\{\mathcal{I}\left[Ve^{j\Phi}e^{j\omega t}\right]\right\} = Ve^{j\Phi} \tag{7.77}$$

The first and last terms in (7.77) are read as "the phasor transform of $v_i(t)$ equals $Ve^{j\Phi}$". Note that $v_i(t)$ is not equal to $Ve^{j\Phi}$ as can be seen from the fact that $v_i(t)$ is a function of time while $Ve^{j\Phi}$ is not. Phasor transforms will be denoted with **bold** letters that are underlined as in $\wp[v_i(t)] = \underline{\mathbf{V}}_i$.

If our analysis is confined to single-frequency cosine waves, we perform the phasor transform in the following manner:

$$\wp\left[V\cos(\omega t + \Phi)\right] = \wp\left\{\mathcal{R}\left[Ve^{j\Phi}e^{j\omega t}\right]\right\} = Ve^{j\Phi} = \underline{\mathbf{V}} \tag{7.78}$$

In other words, to perform the phasor transform of a cosine function, we drop both the real sign and the term $e^{j\omega t}$. Both sines and cosines are sinusoidal functions, but when we transform them, they lose their identities. Thus, before starting an analysis, we must decide whether to perform the analysis all in sines or all in cosines. The two functions must not be mixed when using phasor transforms. Furthermore, we cannot simultaneously employ the phasor transforms of sinusoids at two different frequencies. However, if a linear network has two excitations which have different frequencies, we can use superposition in an analysis for a voltage or current, and add the solutions in the time domain.

Three equivalent representations are used for a phasor $\underline{\mathbf{V}}$:

$$\underline{\mathbf{V}} = \begin{cases} Ve^{j\Phi} & \text{exponential form} \\ V(\cos\Phi + j\sin\Phi) & \text{rectangular form} \\ V\angle\Phi & \text{polar form} \end{cases} \tag{7.79}$$

If phasors are to be multiplied or divided by a complex number, the exponential or polar forms are the most convenient. If phasors are to be added or subtracted, the rectangular form is the most convenient. The relationships among the equivalent representations are illustrated in Figure 7.64. In this figure, the phasor $\underline{\mathbf{V}}$ is denoted by a point in the complex plane. The magnitude of the phasor, $|\underline{\mathbf{V}}| = V$, is illustrated

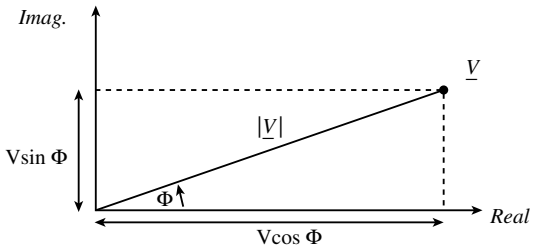


FIGURE 7.64 Relationships among phasor representations.

by the length of the line drawn from the origin to the point. The phase of the phasor, Φ, is shown measured counterclockwise from the horizontal axis. The real part of $\underline{\boldsymbol{V}}$ is $V\cos\Phi$ and the imaginary part of $\underline{\boldsymbol{V}}$ is $V\sin\Phi$.

Phasors can be developed in a way that parallels, to some extent, the usual development of Laplace transforms. In the following theorems, we assume that the constants V_1, V_2,Φ_1, and Φ_2 are real.

Theorem 1: *For sinusoids of the same type* (*either sines or cosines*) *and of the same frequency* ω, $\wp[V_1\sin(\omega t + \Phi_1) + V_2\sin(\omega t + \Phi_2)] = V_1\wp[\sin(\omega t + \Phi_1)] + V_2\wp[\sin(\omega t + \Phi_2)]$. *A similar relation can be written for cosines.*

This theorem demonstrates that the phasor transform is a linear transform.

Theorem 2: *If* $\wp[V_1\sin(\omega t + \Phi)] = V_1e^{j\Phi}$, *then*

$$\wp\left[\frac{d}{dt}V_1\sin(\omega t+\Phi)\right] = j\omega V_1e^{j\Phi} \tag{7.80}$$

To prove Theorem 2, we can write:

$$\wp\left[\frac{d}{dt}V_1\mathcal{I}\left(e^{j\Phi}e^{j\omega t}\right)\right] = \wp\left[V_1\mathcal{I}\left(e^{j\Phi}\frac{d}{dt}e^{j\omega t}\right)\right]$$
$$= \wp\left[V_1\mathcal{I}e^{j\Phi}j\omega e^{j\omega t}\right] = V_1j\omega e^{j\Phi}$$

Note the interchange of the derivative and the imaginary sign in the proof of the theorem. Also, Theorem 2 can be generalized to:

$$\wp\left[\frac{d^n}{dt^n}V_1\sin(\omega t+\Phi)\right] = (j\omega)^nV_1e^{j\Phi} \tag{7.81}$$

These results are useful for finding the sinusoidal steady state solutions of linear, constant-coefficient, ordinary differential equations assuming the roots of the characteristic polynomials lie in the open LHP with possibly one at the origin.

Theorem 3: *If* $\wp[V_1\sin(\omega t + \Phi)] = V_1e^{j\Phi}$, *then*

$$\wp\left[\int V_1\sin(\omega t+\Phi)dt\right] = \frac{1}{j\omega}V_1e^{j\Phi} \tag{7.82}$$

The proof of Theorem 3 is easily obtained by writing:

$$\wp\left[\int V_1 \sin(\omega t + \Phi)\, dt\right] = \wp\left[\int \Im\left[V_1 e^{j(\omega t+\Phi)}\right] dt\right]$$

$$= \wp\left[\Im \int V_1 e^{j(\omega t+\Phi)}\, dt\right] = \frac{V_1}{j\omega} e^{j\Phi}$$

It should be noted that no constant of integration is employed because a constant is not a sinusoidal function and is therefore not permitted when using phasors. A constant of integration arises in LLFT network analysis because of initial conditions, and we are interested only in the sinusoidal steady-state response and not in a zero-input response. No limits are used with the integral either, because the (constant) lower limit would also yield a constant, which would imply that we are not at sinusoidal steady state.

Theorem 3 is easily extended to the case of *n* integrals:

$$\wp\left[\int \ldots \int V_1 \sin(\omega t + \Phi)(dt)^n\right] = \frac{V_1}{(j\omega)^n} e^{j\Phi} \tag{7.83}$$

This result is useful for finding the sinusoidal steady-state solution of integro-differential equations.

Inverse Phasor Transforms

To obtain time domain results, we must be able to inverse transform phasors. The inverse transform operation is denoted by $\wp^{-1}$. This is an easy operation that consists of restoring the term $e^{j\omega t}$, restoring the imaginary sign (the real sign if cosines are used), and dropping the inverse transform sign. That is,

$$\begin{aligned} \wp^{-1}\left[V_1 e^{j\Phi}\right] &= \Im\left[V_1 e^{j\Phi} e^{j\omega t}\right] \\ &= V_1 \sin(\omega t + \Phi) \end{aligned} \tag{7.84}$$

The following example illustrates both the use of Theorem 2 and the inverse transform procedure.

Example 23. Determine the sinusoidal steady-state solution for the differential equation:

$$\frac{d^2 f(t)}{dt^2} + 4\frac{df(t)}{dt} + 3f(t) = V \sin(\omega t + \Phi)$$

Solution. We note that the characteristic polynomial, $D^2 + 4D + 3$, has all its roots in the open LHP. The next step is to phasor transform each term of the equation to obtain:

$$-\omega^2 \underline{\boldsymbol{F}} + 4j\omega \underline{\boldsymbol{F}} + 3\underline{\boldsymbol{F}} = Ve^{j\Phi}$$

where $\underline{\boldsymbol{F}}(j\omega) = \wp[f(t)]$. Therefore, when we solve for $\underline{\boldsymbol{F}}$, we obtain

$$\underline{\boldsymbol{F}} = \frac{Ve^{j\Phi}}{(3-\omega^2) + j4\omega}$$

$$= \frac{Ve^{j\Phi}e^{j\theta}}{\sqrt{(3-\omega^2)^2 + 16\omega^2}}$$

where

$$\theta = \tan^{-1}\frac{-4\omega}{3-\omega^2} = \tan^{-1}\frac{4\omega}{\omega^2 - 3}$$

Thus,

$$\underline{\boldsymbol{F}} = \frac{V}{\sqrt{\omega^4 + 10\omega^2 + 9}} e^{j(\phi+\theta)}$$

To obtain a time-domain function, we inverse transform $\underline{\boldsymbol{F}}$ to obtain:

$$\mathcal{F}^{-1}\left[\underline{\boldsymbol{F}}(j\omega)\right] = f(t) = \frac{V}{\sqrt{\omega^4 + 10\omega^2 + 9}} \sin(\omega t + \Phi + \theta)$$

In this example, we see that the sinusoidal steady-state solution consists of the sinusoidal forcing term, $V \sin(\omega t + \Phi)$, modified in amplitude and shifted in phase.

Phasors and Networks

Phasors are time-independent representations of sinusoids. Thus, we can define impedances in the phasor transform domain and obtain Ohm's law-like expressions relating currents through network elements with the voltages across those elements. In addition, the impedance concept allows us to combine dissimilar elements, such as resistors with inductors, in the transform domain.

The time-domain expressions relating the voltages and currents for *Rs*, *Ls*, and *Cs*, repeated here for convenience, are:

$$v_R(t) = i_R(t)R \qquad v_L(t) = \frac{Ldi_L}{dt} \qquad v_c(t) = \frac{1}{C}\int i_C dt$$

Note that initial conditions are set to zero. Then, performing the phasor transform of the time-domain variables, we have

$$Z_R = R \qquad Z_L = j\omega L \qquad Z_C = \frac{1}{j\omega C}$$

We can also write the admittances of these elements as $Y_R = 1/Z_R$, $Y_L = 1/Z_L$, and $Y_C = 1/Z_C$. Then, we can extend the impedance and admittance concepts for two-terminal elements to multiport networks in the same manner as was done in the development of Laplace transform techniques for network analysis. For example, the transfer function of the circuit shown in Figure 7.65 can be written as:

$$\frac{\underline{\boldsymbol{V}}_0(j\omega)}{\underline{\boldsymbol{V}}_i(j\omega)} = G_{21}(j\omega)$$

where the "$j\omega$" indicates that the analysis is being performed at sinusoidal steady state [1]. It is also assumed that no other excitations exist in *N* in Figure 7.65. With impedances and transfer functions defined, then all the theorems developed for Laplace transform analysis, including source transformations, have a phasor transform counterpart.

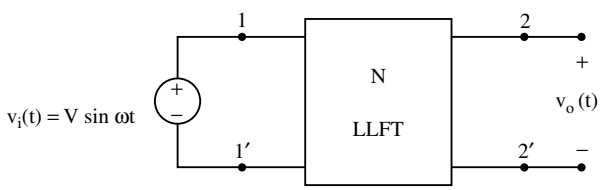


FIGURE 7.65 An LLFT network excited by a sinusoidal voltage source.

Example 21. Use phasor analysis to find the transfer function G_{21} (jω) and v_{oss} (t) for the circuit in Figure 7.66.

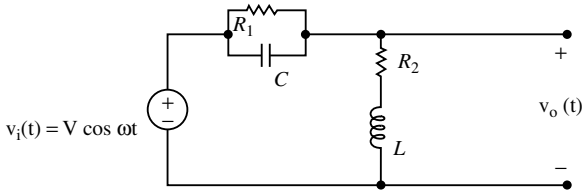


FIGURE 7.66 Circuit for Example 21.

Solution. The phasor transform of the output voltage can be obtained easily by means of the simple voltage divider. Thus,

$$\underline{\boldsymbol{V}_o} = \underline{\boldsymbol{V}_i}\frac{j\omega L + R_2}{j\omega L + R_2 + \dfrac{R_1}{1 + j\omega CR_1}}$$

To obtain $G_{21}(j\omega)$, we form $\underline{\boldsymbol{V}}_0/\underline{\boldsymbol{V}}_i$, which yields

$$\frac{\underline{\boldsymbol{V}_0}}{\underline{\boldsymbol{V}_i}} = G_{21}(j\omega) = \frac{(R_2 + j\omega L)(1 + j\omega CR_1)}{(R_2 + j\omega L)(1 + j\omega CR_1) + R_1}$$

Expressing the numerator and denominator of G_{21} in exponential form produces:

$$G_{21} = \frac{\sqrt{\left(R_2 - \omega^2 LCR_1\right)^2 + \left(\omega L + \omega CR_1R_2\right)^2}\, e^{j\alpha}}{\sqrt{\left(R_1 + R_2 - \omega^2 LCR_1\right)^2 + \left(\omega L + \omega CR_1R_2\right)^2}\, e^{j\beta}}$$

where

$$\alpha = \tan^{-1}\frac{\left(\omega L + \omega CR_1R_2\right)}{R_2 - \omega^2 LCR_1}$$

$$\beta = \tan^{-1}\frac{\left(\omega L + \omega CR_1R_2\right)}{R_1 + R_2 - \omega^2 LCR_1}$$

Thus,

$$G_{21}(j\omega) = Me^{j\theta}$$

where

$$M = \sqrt{\frac{\left(R_2 - \omega^2 LCR_1\right)^2 + \left(\omega L + \omega CR_1R_2\right)^2}{\left(R_1 + R_2 - \omega^2 LCR_1\right)^2 + \left(\omega L + \omega CR_1R_2\right)^2}}$$

and

$$\theta = \alpha - \beta$$

The phasor transform of $v_i(t)$ is

$$\underline{\boldsymbol{V}}_i = \mathcal{P}\left[V\mathcal{R}e^{j\omega t}\right] = Ve^{j0} = V$$

and, therefore, the time-domain expression for the sinusoidal steady-state output voltage is:

$$v_{oss}(t) = VM\cos(\omega t + \theta)$$ □

Driving point impedances and admittances as well as transfer functions are not phasors because they do not represent sinusoidal waveforms. However, an impedance or transfer function is a complex number at a particular real frequency, and the product of a complex number times a phasor is a new phasor.

The product of two arbitrary phasors is not ordinarily defined because $\sin^2(\omega t)$ or $\cos^2(\omega t)$ are not sinusoidal and have no phasor transforms. However, as we will see later, power relations for AC circuits can be expressed in efficient ways as functions of products of phasors. Because such products have physical interpretations, we permit them in the context of power calculations.

Division of one phasor by another is permitted only if the two phasors are related by a driving point or transfer network function such as $\underline{\boldsymbol{V}}_0/\underline{\boldsymbol{V}}_i = G_{21}(j\omega)$.

Phase Lead and Phase Lag

The terms "phase lead" and "phase lag" are used to describe the phase shift between two or more sinusoids of the same frequency. This phase shift can be expressed as an angle in degrees or radians, or it can be expressed in time as seconds. For example, suppose we have three sinusoids given by:

$$v_1(t) = V_1\sin(\omega t) \qquad v_2(t) = V_2\sin(\omega t + \Phi) \qquad v_3 = V_3\sin(\omega t - \Phi)$$

where V_1, V_2, V_3, and Φ are all positive. Then, we say that v_2 **leads** v_1 and that v_3 **lags** v_1. To see this more clearly, we rewrite v_2 and v_3 as:

$$v_2 = V_2\sin\left[\omega(t + t_0)\right] \qquad\qquad v_3 = V_3\sin\left[\omega(t - t_0)\right]$$

where the constant $t_0 = \Phi/\omega$. Figure 7.67 plots the three sinusoids sketched on the same axis, and from this graph we see that the zero crossings of $v_2(t)$ occur t_0 seconds before the zero crossing of $v_1(t)$. Thus, $v_2(t)$ leads $v_1(t)$ by t_0 seconds. Similarly, we see that the zero crossings of $v_3(t)$ occur t_0 seconds after the zero crossings of $v_1(t)$. Thus, $v_3(t)$ lags $v_1(t)$. We can also say that $v_3(t)$ lags $v_2(t)$. When comparing the phases of sine waves with $V\sin(\omega t)$, the key thing to look for in the arguments of the sines are the signs of the angles following ωt. A positive sign means lead and a negative sign means lag. If two sines or two cosines have the same phase angle, then they are called "in phase."

If we have $i_1(t) = I_1\,[\cos(\omega t - \pi/4)]$ and $i_2(t) = I_2\,[\cos(\omega t - \pi/3)]$, then i_2 lags i_1 by $\pi/12$ rad or 15° because even though the phases of both cosines are negative, the phase of $i_1(t)$ is less negative than the phase of $i_2(t)$. We can also say that i_1 leads i_2 by 15°.

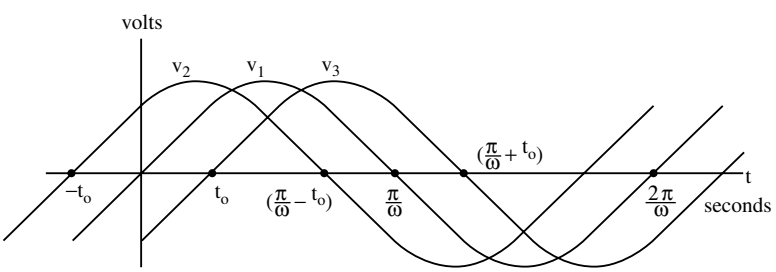


FIGURE 7.67 Three sinusoids sketched on a time axis.

Example 22. Suppose we have five signals with equal peak amplitudes and equal frequencies but with differing phases. The signals are: $i_1 = I\,[\sin(\omega t)]$, $i_2 = I\,[\cos(\omega t)]$, $i_3 = I\,[\cos(\omega t + \theta)]$, $i_4 = -I\,[\sin(\omega t + \psi)]$, and $i_5 = -I\,[\cos(\omega t - \Phi)]$. Assume I, θ, ψ, and Φ are positive.

A. How much do the signals i_2 through i_5 lead i_1?
B. How much do the signals i_1 and i_3 through i_5 lead i_2?

Solution. For part (A), we express i_2 through i_5 as sines with lead angles. That is,

$$i_2 = I\cos(\omega t) = I\sin\left(\omega t + \frac{\pi}{2}\right)$$

$$i_3 = I\cos(\omega t + \theta) = I\sin\left(\omega t + \theta + \frac{\pi}{2}\right)$$

$$i_4 = -I\sin(\omega t + \psi) = I\sin(\omega t + \psi \pm \pi)$$

$$i_5 = -I\cos(\omega t - \Phi) = I\cos(\omega t - \Phi \pm \pi)$$

$$= I\sin\left(\omega t - \Phi \pm \pi + \frac{\pi}{2}\right)$$

Thus, i_2 leads i_1 by $\pi/2$ rad, and i_3 leads i_1 by $\theta + \pi/2$. For i_4, we can take the plus sign in the argument of the sign to obtain $\psi + \pi$, or we can take the minus sign to obtain $\psi - \pi$. The current i_5 leads i_1 by $(3\pi/2 - \Phi)$ or by $(-\pi/2 - \Phi)$. An angle of $\pm 2\pi$ can be added to the argument without affecting lead or lag relationships.

For part (B), we express i_1 and i_3 through i_5 as cosines with lead angles yielding:

$$i_1 = I\sin(\omega t) = I\cos\left(\omega t - \frac{\pi}{2}\right)$$

$$i_3 = I\cos(\omega t + \theta)$$

$$i_4 = -I\sin(\omega t + \psi) = I\sin(\omega t + \psi \pm \pi)$$

$$= I\cos\left(\omega t + \psi \pm \pi - \frac{\pi}{2}\right)$$

$$i_5 = -I\cos(\omega t - \Phi) = I\cos(\omega t - \Phi \pm \pi)$$

We conclude that i_1 leads i_2 by $(-\pi/2)$ rad. (We could also say that i_1 lags i_2 by $(\pi/2)$ rad.) Also, i_3 leads i_2 by θ. The current i_4 leads i_2 by $(\psi + \pi/2)$ where we have chosen the plus sign in the argument of the cosine. Finally, i_5 leads i_2 by $(\pi - \Phi)$, where we have chosen the plus sign in the argument. □

In the previous example, we have made use of the identities:

$$\cos(\alpha) = \sin\left(\alpha + \frac{\pi}{2}\right); \qquad -\sin(\alpha) = \sin(\alpha \pm \pi)$$

$$-\cos(\alpha) = \cos(\alpha \pm \pi); \qquad \sin(\alpha) = \cos\left(\alpha - \frac{\pi}{2}\right)$$

The concepts of phase lead and phase lag are clearly illustrated by means of phasor diagrams, which are described in the next section.

Phasor Diagrams

Phasors are complex numbers that represent sinusoids, so phasors can be depicted graphically on a complex plane. Such graphical illustrations are called phasor diagrams. Phasor diagrams are valuable because they present a clear picture of the relationships among the currents and voltages in a network. Furthermore, addition and subtraction of phasors can be performed graphically on a phasor diagram. The construction of phasor diagrams is demonstrated in the next example.

Example 23. For the network in Figure 7.68(a), find $\underline{\boldsymbol{I}}_1$, $\underline{\boldsymbol{V}}_{R1}$, and $\underline{\boldsymbol{V}}_C$. For Figure 7.68(b), find $\underline{\boldsymbol{I}}_2$, $\underline{\boldsymbol{V}}_{R2}$, and $\underline{\boldsymbol{V}}_L$. Construct phasor diagrams that illustrate the relations of the currents to the voltage excitation and the other voltages of the networks.

Solution. For Figure 7.68(a), we have

$$\mathcal{P}\left[v(t)\right] = V\angle 0^\circ \quad \text{and} \quad \mathcal{P}\left[i_1(t)\right] = \underline{\boldsymbol{I}}_1 = \frac{V}{R_1 + \dfrac{1}{j\omega C}}$$

Rewriting $\underline{\boldsymbol{I}}_1$, we have:

$$\underline{\boldsymbol{I}}_1 = \frac{Vj\omega C}{1 + j\omega CR_1} = \frac{Vj\omega C}{1 + j\omega CR_1}\left[\frac{1 - j\omega CR_1}{1 - j\omega CR_1}\right]$$

$$= V\left[\frac{\omega^2 C^2 R_1 + j\omega C}{\omega^2 C^2 R_1^2 + 1}\right] = \frac{V\omega C}{\sqrt{\omega^2 C^2 R_1^2 + 1}} e^{j\theta_1}$$

where

$$\theta_1 = \tan^{-1}\frac{\omega C}{\omega^2 C^2 R_1} = \tan^{-1}\frac{1}{\omega CR_1}$$

Note that we have multiplied the numerator and denominator of $\underline{\boldsymbol{I}}_1$ by the conjugate of the denominator. The resulting denominator of $\underline{\boldsymbol{I}}_1$ is purely real, and so we need only consider the terms in the numerator of $\underline{\boldsymbol{I}}_1$ to obtain an expression for the phase. Thus, the resulting expression for the phase contains only one term which has the form:

$$\theta_1 = \tan^{-1}\frac{\mathcal{I}(\text{numerator})}{\mathcal{R}(\text{numerator})}$$

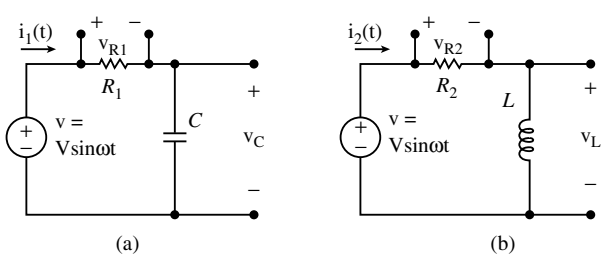


FIGURE 7.68 (a) An RC network. (b) An RL network.

We could have obtained the same results without application of this artifice. In this case, we would have obtained

$$\theta_1 = \frac{\pi}{2} - \tan^{-1}\omega CR_1$$

For $\omega CR_1 \geq 0$, it is easy to show that the two expressions for θ_1 are equivalent.

Because the same current flows through both network elements, we have

$$\underline{\boldsymbol{V}}_{R1} = \frac{V\omega CR_1}{\sqrt{\omega^2 C^2 R_1^2 + 1}} e^{j\theta_1}$$

and

$$\underline{\boldsymbol{V}}_c = \underline{\boldsymbol{I}}_1\left(\frac{1}{j\omega C}\right) = \frac{-jV}{\sqrt{\omega^2 C^2 R_1^2 + 1}} e^{j\theta_1} = \frac{V}{\sqrt{\omega^2 C^2 R_1^2 + 1}} e^{j\psi}$$

where

$$\psi = -\frac{\pi}{2} + \theta_1 = -\tan^{-1}\omega CR_1$$

For $\underline{\boldsymbol{I}}_2$ in Figure 7.68(b), we obtain

$$\underline{\boldsymbol{I}}_2 = \frac{V\angle 0^o}{R_2 + j\omega L} = \frac{V}{\sqrt{R_2^2 + w^2 L^2}} e^{j\theta_2}$$

where θ_2 is given by

$$\theta_2 = -\tan^{-1}\frac{\omega L}{R_2}$$

The phasor current $\underline{\boldsymbol{I}}_2$ flows through both R_2 and L. So we have:

$$\underline{\boldsymbol{V}}_{R2} = \underline{\boldsymbol{I}}_2 R_2$$

and

$$\underline{\boldsymbol{V}}_L = j\omega L\,\underline{\boldsymbol{I}}_2 = \frac{V\omega L}{\sqrt{\omega^2 L^2 + R_2^2}} e^{j\Phi}$$

where

$$\Phi = \frac{\pi}{2} + \theta_2$$

To construct the phasor diagram in Figure 7.69(a) for the RC network in Figure 7.68(a), we first draw a vector corresponding to the phasor transform $\underline{\boldsymbol{V}} = V\angle 0°$ of the excitation. Because the phase of this phasor is zero, it is represented as a vector along the positive real axis. The length of this vector is $|\underline{\boldsymbol{V}}|$. Then we construct the vector representing $\underline{\boldsymbol{I}}_1 = |\underline{\boldsymbol{I}}_1| e^{j\theta 1}$. Again, the length of the vector is $|\underline{\boldsymbol{I}}_1|$, and it is

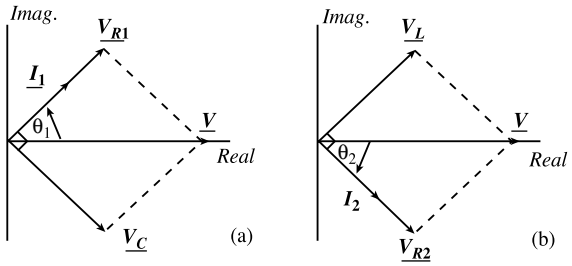


FIGURE 7.69 (a) Phasor diagram for the voltages and currents in Figure 7.68(a). (b) Phasor diagram for Figure 7.68(b).

drawn at the angle θ_1. The vector representing $\underline{\boldsymbol{V}}_{R1}$ lies along $\underline{\boldsymbol{I}}_1$ because the voltage across a resistor is always in phase or 180° out of phase with the current flowing through the resistor. The vector representing the current leads $\underline{\boldsymbol{V}}_C$ by exactly 90°. It should be noted from the phasor diagram that $\underline{\boldsymbol{V}}_{R1}$ and $\underline{\boldsymbol{V}}_C$ add to produce $\underline{\boldsymbol{V}}$ as required by Kirchhoff's law. □

Figure 7.69(b) presents the phasor diagram for the RL network in Figure 7.68(b). For this network, $\underline{\boldsymbol{I}}_2$ lags $\underline{\boldsymbol{V}}_L$ by exactly 90°. Also, the vector sum of the voltages $\underline{\boldsymbol{V}}_L$ and $\underline{\boldsymbol{V}}_{R2}$ must be the excitation voltage $\underline{\boldsymbol{V}}$ as indicated by the dotted lines in Figure 7.69(b).

If the excitation $V\sin(\omega t)$ had been $V\sin(\omega t + \Phi)$ in Figure 7.68 in the previous example, then the vectors in the phasor diagrams in Figure 7.69 would have just been rotated around the origin by Φ. Thus, for example, $\underline{\boldsymbol{I}}_1$ in Figure 7.69(a) would have an angle equal to $\theta_1 + \Phi$. The lengths of the vectors and the relative phase shifts between the vectors would remain the same.

If R_1 in Figure 7.68(a) is decreased, then from the expression for $\theta_1 = \tan^{-1}(1/(\omega CR_1))$, we see that the phase of $\underline{\boldsymbol{I}}_1$ is increased. As R_1 is reduced further, θ_1 approaches 90°, and the circuit becomes more nearly like a pure capacitor. However, as long as $\underline{\boldsymbol{I}}_1$ leads $\underline{\boldsymbol{V}}$, we label the circuit as capacitive.

As R_2 in Figure 7.68(b) is decreased, then θ_2 in Figure 7.69(b) decreases (becomes more negative) and approaches –90°. Nevertheless, as long as $\underline{\boldsymbol{I}}_2$ lags $\underline{\boldsymbol{V}}$, we refer to the circuit as inductive.

If both inductors and capacitors are in a circuit, then it is possible for the circuit to appear capacitive at some frequencies and inductive at others. An example of such a circuit is provided in the next section.

Resonance

Resonant networks come in two basic types: the parallel resonant network and the series resonant (sometimes called antiresonant) network. More complicated networks may contain a variety of both types of resonant circuits. To see what happens at resonance, we examine a parallel resonant network at sinusoidal steady state [1]. Figure 7.70 is a network consisting of a capacitor and inductor connected in parallel, often called a tank circuit or tank, and an additional resistor R_1 connected in parallel with the tank. The phasor transforms of the excitation and the currents through the elements in Figure 7.70 are:

$$\overline{V} = V\angle 0^\circ;\ \underline{\boldsymbol{I}}_{R1} = \frac{V}{R_1};\ \underline{\boldsymbol{I}}_C = j\omega CV;\ \underline{\boldsymbol{I}}_L = \frac{V}{j\omega L} \tag{7.85}$$

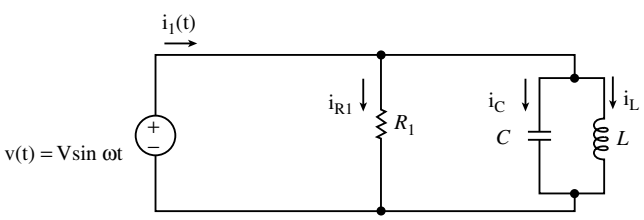


FIGURE 7.70 Parallel resonant circuit.

where V is the peak value of the excitation. The transform of the current supplied by the source is

$$\underline{\boldsymbol{I}}_1 = \boldsymbol{I}_1 \angle \theta_1 = \underline{\boldsymbol{I}}_{R1} + \underline{\boldsymbol{I}}_C + \underline{\boldsymbol{I}}_L = V\left[\frac{1}{R_1} + j\omega C\left(1 - \frac{1}{\omega^2 LC}\right)\right] \tag{7.86}$$

The peak value of the current $i_1(t)$ at steady state is

$$I_1 = V\sqrt{\left(\frac{1}{R_1}\right)^2 + \omega^2 C^2\left(1 - \frac{1}{\omega^2 LC}\right)^2} \tag{7.87}$$

It is not difficult to determine that the minimum value of I_1 occurs at

$$\omega = \frac{1}{\sqrt{LC}} \tag{7.88}$$

which is the condition for resonance, and $I_{1\min}$ is given by

$$I_{1\min} = \frac{V}{R_1} \tag{7.89}$$

This result is somewhat surprising since it means that at resonance, the source in Figure 7.70 delivers no current to the tank at steady state. However, this result does not mean that the currents through the capacitor and inductor are zero. In fact, for $\omega^2 = 1/(LC)$ we have:

$$\underline{\boldsymbol{I}}_C = jV\sqrt{\frac{C}{L}} \text{ and } \underline{\boldsymbol{I}}_L = -jV\sqrt{\frac{C}{L}}$$

That is, the current through the inductor is 180° out of phase with the current through the capacitor, and, because their magnitudes are equal, their sum is zero. Thus, at steady state and at the frequency given by (7.88), the tank circuit looks like an open circuit to the voltage source. Yet, a circulating current occurs in the tank, labeled $\underline{\boldsymbol{I}}_T$ in Figure 7.71, which can be quite large depending on the values of C and L. That is, at resonance,

$$\underline{\boldsymbol{I}}_T = jV\sqrt{\frac{C}{L}} = \underline{\boldsymbol{I}}_C = -\underline{\boldsymbol{I}}_L \tag{7.90}$$

Therefore, energy is being transferred back and forth between the inductor and the capacitor. If the inductor and capacitor are ideal, the energy transferred would never decrease. In practice, parasitic resistances, especially in a physical inductor, would eventually dissipate this energy. Of course, parasitic resistances can be modeled as additional elements in the network.

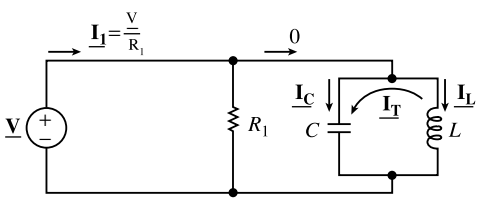


FIGURE 7.71 Circuit of Figure 7.70 at resonance. No current is supplied to the tank by the source, but a circulating current occurs in the tank.

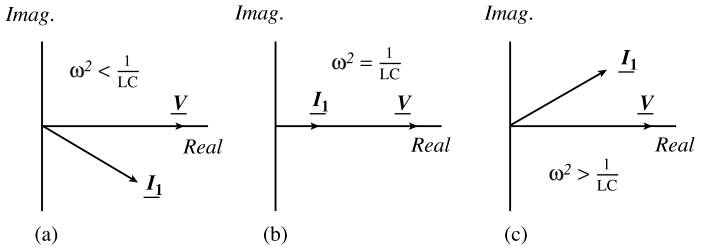


FIGURE 7.72 Phasor diagrams for the circuit in Figure 7.70. (a) $\omega^2 < 1/LC$. (b) Diagram at resonance. (c) $\omega^2 > 1/(LC)$.

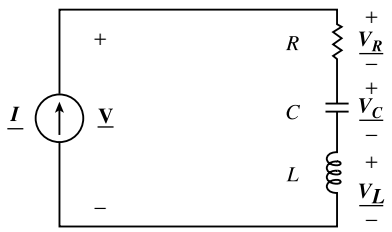


FIGURE 7.73 Series resonant circuit.

Another interesting aspect of the network in Figure 7.70 is that, at low frequencies ($\omega^2 < 1/(LC)$), $\underline{\boldsymbol{I}}_1$ lags $\underline{\boldsymbol{V}}$, and so the network appears inductive to the voltage source. At high frequencies ($\omega^2 > 1/(LC)$), $\underline{\boldsymbol{I}}_1$ leads $\underline{\boldsymbol{V}}$, and the network looks capacitive to the voltage source. At resonance, the network appears as only a resistor R_1 to the source. Figure 7.72 depicts phasor diagrams of $\underline{\boldsymbol{V}}$ and $\underline{\boldsymbol{I}}_1$ at low frequency, at resonance, and at high frequency.

Figure 7.73 is the second basic type of resonant circuit — a series resonant circuit which is excited by a sinusoidal current source with phasor transform $\underline{\boldsymbol{I}} = I \angle 0°$. This circuit is dual to the circuit in Figure 7.70. The voltages across the network elements can be expressed as:

$$\underline{\boldsymbol{V}}_R = IR;\ \underline{\boldsymbol{V}}_C = -j\left(\frac{1}{\omega C}\right)I;\ \underline{\boldsymbol{V}}_L = j\omega LI \tag{7.91}$$

Then, the voltage $\underline{\boldsymbol{V}}$ is

$$\underline{\boldsymbol{V}} = I\left[R + j\left(\omega L - \frac{1}{\omega C}\right)\right] \tag{7.92}$$

The peak value of $\underline{\boldsymbol{V}}$ is

$$V = I\sqrt{R^2 + \left(\omega L - \frac{1}{\omega C}\right)^2} \tag{7.93}$$

where I is the peak value of $\underline{\boldsymbol{I}}$. The minimum value of V is

$$V_{\min} = IR \tag{7.94}$$

and this occurs at the frequency $\omega = 1/\sqrt{(LC)}$, which is the same resonance condition as for the circuit in Figure 7.70.

Equation (7.94) demonstrates that at resonance, the voltage across the LC subcircuit in Figure 7.73 is zero. However, the individual voltages across L and across C are not zero and can be quite large in magnitude depending on the values of the capacitor and inductor. These voltages are given by:

$$\underline{V_C} = -jI\sqrt{\frac{L}{C}} \quad \text{and} \quad \underline{V_L} = jI\sqrt{\frac{L}{C}} \tag{7.95}$$

and therefore the voltage across the capacitor is exactly 180° out of phase with the voltage across the inductor.

At frequencies below resonance, $\underline{\boldsymbol{V}}$ lags $\underline{\boldsymbol{I}}$ in Figure 7.73, and therefore the circuit looks capacitive to the source. Above resonance, $\underline{\boldsymbol{V}}$ leads $\underline{\boldsymbol{I}}$, and the circuit looks inductive to the source. If the frequency of the source is $\omega = 1/\sqrt{(LC)}$ the circuit looks like a resistor of value R to the source.

Power in AC Circuits

If a sinusoidal voltage $v(t) = V\sin(\omega t + \theta_V)$ is applied to an LLFT network that possibly contains other sinusoidal sources having the same frequency ω, then a sinusoidal current $i(t) = I\sin(\omega t + \theta_1)$ flows at steady state as depicted in Figure 7.74. The instantaneous power delivered to the circuit by the voltage source is

$$p(t) = v(t)i(t) = VI\sin(\omega t + \theta_v)\sin(\omega t + \theta_I) \tag{7.96}$$

where the units of $p(t)$ are watts (W). With the aid of the trigonometric identity

$$\sin\alpha\sin\beta = \frac{1}{2}\left[\cos(\alpha - \beta) - \cos(\alpha + \beta)\right]$$

we rewrite (7.96) as

$$p(t) = \frac{1}{2}VI\left[\cos(\theta_v - \theta_I) - \cos(2\omega t + \theta_v + \theta_I)\right] \tag{7.97}$$

The instantaneous power delivered to the network in Figure 7.74 has a component that is constant and another component that has a frequency twice that of the excitation. At different instances of time, $p(t)$ can be positive or negative, meaning that the voltage source is delivering power to the network or receiving power from the network, respectively.

In AC circuits, however, it is usually the average power P that is of more interest than the instantaneous power $p(t)$ because average power generates the heat or performs the work.

The average over a period of a periodic function $f(t)$ with period T is

$$\left[f(t)\right]_{\text{avg}} = F = \frac{1}{T}\int_0^T f(t)dt \tag{7.98}$$

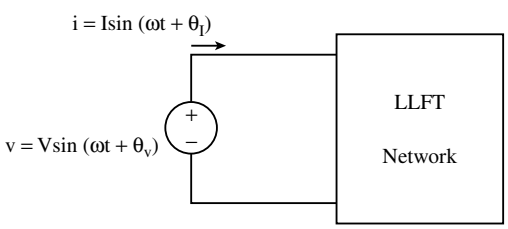


FIGURE 7.74 LLFT network that may contain other sinusoidal sources at the same frequency as the external generator.

The period of $p(t)$ in (7.97) is $T = \pi/\omega$, and so

$$\left[p(t)\right]_{\text{avg}} = P = \frac{\omega}{\pi}\int_0^{\frac{\pi}{\omega}} p(t)dt = \frac{1}{2}VI\cos\left(\theta_V - \theta_I\right) \tag{7.99}$$

The cosine term in (7.99) plays an important role in power calculations and so is designated as the Power Factor (*PF*). Thus,

$$\text{Power Factor} = PF = \cos\left(\theta_V - \theta_I\right) \tag{7.100}$$

If $|\theta_V - \theta_I| = \pi/2$, then $PF = 0$, and the average power delivered to the network in Figure 7.74 is zero; but if $PF = 1$, then P delivered to the network by the source is $VI/2$. If $0 < |\theta_v - \theta_I| < \pi/2$, then P is positive, and the source is delivering average power to the network. However, the network delivers average power to the source when P is negative, and this occurs if $\pi/2 < |\theta_v - \theta_I| < 3\pi/2$.

If the current leads the voltage in Figure 7.74, the convention is to consider *PF* as leading, and if current lags the voltage, the *PF* is regarded as lagging. However, it is not possible from *PF* alone to determine whether a current leads or lags voltage.

Example 24. Determine the average power delivered to the network shown in Figure 7.68(a).

Solution. The phasor transform of the applied voltage is $\underline{\boldsymbol{V}} = V\angle 0°$, and we determined in Example 23 that the current supplied was

$$\underline{\boldsymbol{I}}_1 = \frac{V\omega C e^{j\theta_1}}{\sqrt{\omega^2 C^2 R_1^2 + 1}}, \qquad \theta_1 = \tan^{-1}\frac{1}{\omega C R_1}$$

The power factor is

$$PF = \cos\left(0 - \theta_1\right) = \cos\left(\theta_1\right)$$

which, with the aid of the triangle in Figure 7.75, can be rewritten as

$$PF = \frac{\omega C R_1}{\sqrt{\left(\omega C R_1\right)^2 + 1}}$$

Thus, the average power delivered to the circuit is

$$P = \frac{1}{2}\frac{V^2\omega C}{\sqrt{\left(\omega C R_1\right)^2 + 1}}\left[\frac{\omega C R_1}{\sqrt{\left(\omega C R_1\right)^2 + 1}}\right] = \frac{V^2\omega^2 C^2 R_1}{2\left(\omega^2 C^2 R_1^2 + 1\right)} = \frac{I_1^2 R_1}{2}$$

□

We note that if R_1 were zero in the previous example, then $P = 0$ because the circuit would be purely capacitive, and *PF* would be zero.

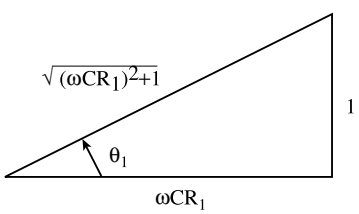


FIGURE 7.75 Triangle for determining *PF*.

If no sources are in the network in Figure 7.74, then the network terminal variables are related by:

$$\underline{\boldsymbol{V}} = \underline{\boldsymbol{I}}\, Z(j\omega) \tag{7.101}$$

where $Z(j\omega)$ is the input impedance of the network. Because Z is, general and complex, we can write it as:

$$Z(j\omega) = R(\omega) + jX(\omega) = |Z|e^{j\theta_Z} \tag{7.102}$$

where

$$R(\omega) = \Re Z(j\omega); \quad X(\omega) = \Im Z(j\omega)$$
$$\text{and } \theta_Z = \tan^{-1}\left(\frac{X(\omega)}{R(\omega)}\right) \tag{7.103}$$

In (7.103), the (real) function $X(\omega)$ is termed the reactance. Employing the polar form of the phasors, we can rewrite (7.101) as

$$V\angle\theta_V = I\angle\theta_I|Z|\angle\theta_Z = I|Z|\angle(\theta_I + \theta_Z) \tag{7.104}$$

Equating magnitudes and angles, we obtain

$$V = I|Z| \text{ and } \theta_V = \theta_I + \theta_Z \tag{7.105}$$

Thus, we can express P delivered to the network as

$$P = \frac{1}{2}VI\cos(\theta_V - \theta_I) = \frac{1}{2}I^2|Z|\cos\theta_Z \tag{7.106}$$

But $|Z|\cos(\theta_z) = R(\omega)$ so that

$$P = \frac{1}{2}I^2R(\omega) \tag{7.107}$$

Eq. (7.107) indicates that the real part of the impedance absorbs the power. The imaginary part of the impedance, $X(\omega)$, does not absorb average power. Example 24 in this section provides an illustration of (7.107).

An expression for average power in terms of the input admittance $Y(j\omega) = 1/Z(j\omega)$ can also be obtained. Again, if no sources are within the network, then the terminal variables in Figure 7.74 are related by

$$\underline{\boldsymbol{I}} = \underline{\boldsymbol{V}}\, Y(j\omega) \tag{7.108}$$

The admittance $Y(j\omega)$ can be written as

$$Y(j\omega) = |Y(j\omega)|e^{j\theta_Y} = G(\omega) + jB(\omega) \tag{7.109}$$

where $G(\omega)$ is conductance and $B(\omega)$ is susceptance, and where

$$G(\omega) = \Re Y(j\omega); \quad B(\omega) = \Im Y(j\omega)$$
$$\text{and } \theta_Y = \tan^{-1}\left[\frac{B(\omega)}{G(\omega)}\right] \tag{7.110}$$

Then, average power delivered to the network can be expressed as:

$$P = \frac{1}{2}V^2|Y|\cos\theta_Y = \frac{1}{2}V^2G(\omega) \tag{7.111}$$

If the network contains sinusoidal sources, then (7.99) should be employed to obtain P instead of (7.107) or (7.111).

Consider a resistor R with a voltage $v(t) = V\sin(\omega t)$ across it and therefore a current $i(t) = I\sin(\omega t) = v(t)/R$ through it. The instantaneous power dissipated by the resistor is

$$p(t) = v(t)i(t) = \frac{v^2(t)}{R} = i^2(t)R \tag{7.112}$$

The average power dissipated in R is

$$P = \frac{1}{T}\int_0^T i^2(t)\,R dt = I_{eff}^2 R \tag{7.113}$$

where we have introduced the new constant I_{eff}. From (7.113), we can express I_{eff} as

$$I_{eff} = \sqrt{\frac{1}{T}\int_0^T i^2(t)dt} \tag{7.114}$$

This expression for I_{eff} can be read as "the square root of the mean (average) of the square of $i(t)$" or, more simply, as "the root mean square value of $i(t)$," or, even more succinctly, as "the *RMS* value of $i(t)$." Another designation for this constant is I_{rms}. Equation (7.114) can be extended to any periodic voltage or current.

The *RMS* value of a pure sine wave such as $i(t) = I\sin(\omega t + \theta_1)$ or $v(t) = V\sin(\omega t + \theta_v)$ is

$$I_{rms} = \frac{I}{\sqrt{2}} \text{ or } V_{rms} = \frac{V}{\sqrt{2}} \tag{7.115}$$

where I and V are the peak values of the sine waves. Normally, the voltages and currents listed on the nameplates of power equipment and household appliances are given in terms of *RMS* values instead of peak values. For example, a 120-V, 100-W lightbulb is expected to dissipate 100 W when a voltage $120(\sqrt{2})[\sin(\omega t)]$ is impressed across it. The peak value of this voltage is 170 V.

If we employ *RMS* values, (7.99) can be rewritten as

$$P = V_{rms}I_{rms}PF \tag{7.116}$$

Eq. (7.116) emphasizes the fact that the concept of *RMS* values of voltages and currents was developed in order to simplify the calculation of average power.

Because $PF = \cos(\theta_v - \theta_1)$, we can rewrite (7.116) as

$$\begin{aligned} P &= V_{rms}I_{rms}\cos(\theta_V - \theta_I) = \Re\left[V_{rms}e^{j\theta_V}I_{rms}e^{-j\theta_I}\right] \\ &= \Re\left[\underline{\boldsymbol{V}}\,\underline{\boldsymbol{I}}^*\right] \end{aligned} \tag{7.117}$$

where $\underline{\boldsymbol{I}}^*$ is the conjugate of $\boldsymbol{I}$. If $P = \Re\,[\underline{\boldsymbol{V}}\,\underline{\boldsymbol{I}}^*]$, the question arises as to what the imaginary part of $\underline{\boldsymbol{V}}\,\underline{\boldsymbol{I}}^*$ represents. This question leads naturally to the concept of complex power, denoted by the bold letter $\boldsymbol{S}$, which has the units of volt-amperes (VA). If P represents real power, then we can write

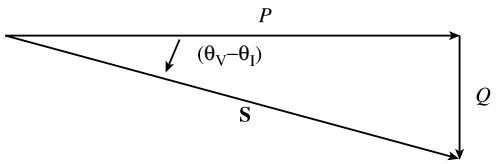


FIGURE 7.76 Power triangle for a capacitive circuit.

$$\boldsymbol{S} = P + jQ \tag{7.118}$$

where

$$\boldsymbol{S} = \underline{\boldsymbol{V}}\,\underline{\boldsymbol{I}}^{*} \tag{7.119}$$

and where

$$Q = \mathcal{I}\left[\underline{\boldsymbol{V}}\,\underline{\boldsymbol{I}}^{*}\right] = V_{rms} I_{rms} \sin\left(\theta_V - \theta_I\right) \tag{7.120}$$

Thus, Q represents imaginary or reactive power. The units of Q are VARs, which stands for volt-amperes reactive. Reactive power is not available for conversion into useful work. It is needed to establish and maintain the electric and magnetic fields associated with capacitors and inductors [4]. It is an overhead required for delivering P to loads, such as electric motors, that have a reactive part in their input impedances.

The components of complex power can be represented on a power triangle. Figure 7.76 is a power triangle for a capacitive circuit. Real and imaginary power are added as shown to yield the complex power **S**. Note that $(\theta_v - \theta_I)$ and Q are both negative for capacitive circuits. The following example illustrates the construction of a power triangle for an RL circuit.

Example 25. Determine the components of power delivered to the RL circuit in Figure 7.77. Provide a phasor diagram for the current and the voltages, construct a power triangle for the circuit, and show how the power diagram is related to the impedances of the circuit.

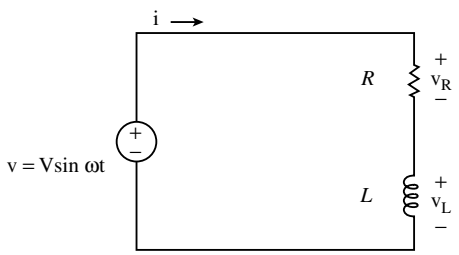


FIGURE 7.77 Network for Example 25.

Solution. We have

$$\underline{\boldsymbol{V}} = Ve^{j0} \text{ and } \underline{\boldsymbol{I}} = \frac{Ve^{j\theta_I}}{\sqrt{R^2 + \left(\omega L\right)^2}}$$

where

$$\theta_I = -\tan^{-1}\frac{\omega L}{R}$$

Because $\theta_v = 0$, *PF* is

$$PF = \cos\left(\theta_V - \theta_I\right) = \frac{R}{\sqrt{R^2 + \left(\omega L\right)^2}}$$

and is lagging. The voltages across *R* and *L* are given by:

$$\underline{\boldsymbol{V}}_{\boldsymbol{R}} = \underline{\boldsymbol{I}} R = \frac{V \operatorname{Re}^{j\theta_I}}{\sqrt{R^2 + \left(\omega L\right)^2}}$$

$$\underline{\boldsymbol{V}}_{\boldsymbol{L}} = j\omega L \underline{\boldsymbol{I}} = \frac{V\omega L}{\sqrt{R^2 + \left(\omega L\right)^2}} e^{j\left(\frac{\pi}{2} + \theta_I\right)}$$

and *Z* is

$$Z = R + j\omega L = \sqrt{R^2 + \left(\omega L\right)^2}\, e^{-j\theta_I}$$

The real and imaginary components of the complex power are simply calculated as:

$$P = I_{rms}^2 R(\omega) = \frac{V_{rms}^2 R}{R^2 + \left(\omega L\right)^2}$$

$$Q = \frac{V_{rms}^2 \omega L}{R^2 + \left(\omega L\right)^2}$$

Figure 7.78 presents the phasor diagram for this circuit in which we have taken the reference phasor as $\underline{\boldsymbol{I}}$ and therefore have shown $\underline{\boldsymbol{V}}$ leading $\underline{\boldsymbol{I}}$ by $(\theta_v - \theta_1)$. Also, we have moved $\underline{\boldsymbol{V}}_L$ parallel to itself to form a triangle. These operations cause the phasor diagram to be similar to the power triangle. Figure 7.79(a) shows a representation for the impedance in Figure 7.77. If each side of the triangle in Figure 7.79(a) is multiplied by I_{rms}, then we obtain voltage triangle in Figure 7.79(b). Next, we multiply the sides of the voltage triangle by I_{rms} again to obtain the power triangle in Figure 7.79(c). The horizontal side is the average power *P*, the vertical side is *Q*, and the hypotenuse has a length that represents the magnitude of the complex power ***S***. All three triangles in Figure 7.79 are similar. The angles between sides are preserved. □

If *P* remains constant in Figure 7.76, but the magnitude of the angle becomes larger so that the magnitude of *Q* increases, then [**S**] increases. If the magnitude of the voltage is fixed, then the magnitude of the current supplied must increase. But then, either power would be lost in the form of heat in the wires supplying the load or larger diameter, more expensive wires, would be needed. For this reason, power companies that supply power to large manufacturing firms that have many large motors impose unfavorable rates. However, the manufacturing firm can improve its rates if it improves its power factor. The following example illustrates how improving (correcting) *PF* is done.

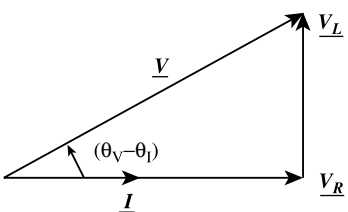


FIGURE 7.78 Phasor diagram for Example 25.

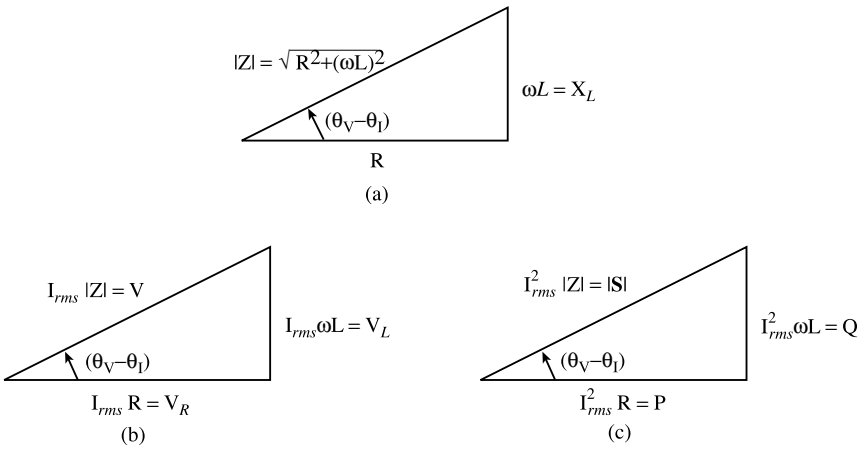


FIGURE 7.79 (a) Impedance triangle for circuit in Example 25. (b) Corresponding voltage triangle. (c) Power triangle.

Example 26. Determine the value of the capacitor to be connected in parallel with the RL circuit in Figure 7.80 to improve the *PF* of the overall circuit to one. The excitation is a voltage source having an amplitude of 120 V *RMS* and frequency $2\pi(60\text{ Hz}) = 377$ rad/s. What are the *RMS* values of the current supplied by the source at steady state before and after the capacitor is connected?

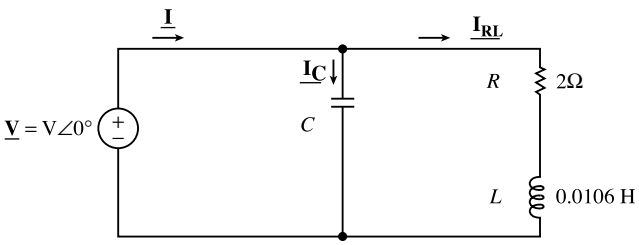


FIGURE 7.80 Circuit for Example 26.

Solution. The current through the RL branch in Figure 7.80 is

$$\underline{\boldsymbol{I}}_{\boldsymbol{RL}} = \frac{Ve^{j\theta}}{\sqrt{R^2 + (\omega L)^2}}; \quad \theta = -\tan^{-1}\frac{\omega L}{R}$$

and the current through the capacitor is

$$\underline{\boldsymbol{I}}_{\boldsymbol{C}} = j\omega CV = V\omega Ce^{j(\pi/2)}$$

Thus, the current supplied by the source to the *RLC* network is

$$\begin{aligned}\underline{\boldsymbol{I}} &= \underline{\boldsymbol{I}}_{\boldsymbol{RL}} + \underline{\boldsymbol{I}}_{\boldsymbol{C}} \\ &= \frac{V\cos\theta}{\sqrt{R^2 + (\omega L)^2}} + jV\left[\frac{-\omega L}{R^2 + (\omega L)^2} + \omega C\right]\end{aligned}$$

To improve the *PF* to one, the current $\underline{\boldsymbol{I}}$ should be in phase with $\underline{\boldsymbol{V}}$. Thus, we set the imaginary term in the equation for $\underline{\boldsymbol{I}}$ equal to zero, yielding:

$$C = \frac{L}{R^2 + (\omega L)^2} = 530\ \mu F$$

a rather large capacitor. Before this capacitor is connected, the *RMS* value of the current supplied by the voltage source is I_{rms} = 26.833 amps. After the capacitor is connected, the source has to supply only 12 amps *RMS*, a considerable reduction. In both cases, *P* delivered to the load is the same.

The following example also illustrates *PF* improvement.

Example 27. A load with *PF* = 0.7 lagging, depicted in Figure 7.81, consumes 12 kW of power. The line voltage supplied is 220 V *RMS* at 60 Hz. Find the size of the capacitor needed to correct the *PF* to 0.9 lagging, and determine the values of the currents supplied by the source both before and after the *PF* is corrected.

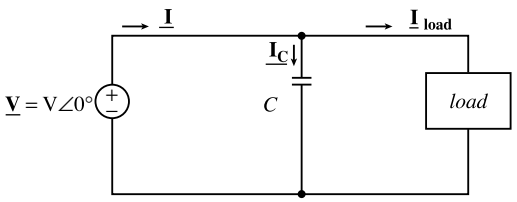


FIGURE 7.81 Circuit for Example 27 showing the load and the capacitor to be connected in parallel with the load to improve the power factor.

Solution. We will take the phase of the line voltage to be 0°. From $P = V_{rms} I_{rms} PF$ = 12 kW, we obtain I_{rms} = 77.922 amps. Because *PF* is 0.7 lagging, the phase of the current through the load relative to the phase of the line voltage is $-\cos^{-1}(0.7) = -45.57°$. Therefore, $\underline{\boldsymbol{I}}_{\boldsymbol{load}} = 77.922\angle(-45.57°)$ amps *RMS*. When *C* is connected in parallel with the load,

$$\underline{\boldsymbol{I}} = \underline{\boldsymbol{I}}_{\boldsymbol{C}} + \underline{\boldsymbol{I}}_{\boldsymbol{load}} = 220(377)jC + 77.922e^{-j0.7954}$$

$$= 54.54 - j[55.64 - 82{,}940C]$$

If the *PF* were to be corrected to unity, we would set the imaginary part of the previous expression for current to zero; but this would require a larger capacitor (671 μF), which may be uneconomical. Instead, to retain a lagging but improved *PF* = 0.9, and corresponding to the current lagging the voltage by 25.84°, we write

$$0.9 = \frac{54.54}{\sqrt{54.54^2 + (55.64 - 82{,}940C)^2}}$$

Therefore, *C* = 352 μF. The line current is now

$$\underline{\boldsymbol{I}} = \underline{\boldsymbol{I}}_{\boldsymbol{C}} + \underline{\boldsymbol{I}}_{\boldsymbol{load}} = 60.615\angle(-25.87^\circ)\text{amps } RMS$$

□

Previous examples have employed ideal voltage sources to supply power to networks. However, in many electronic applications, the source has a fixed impedance associated with it, and the problem is to

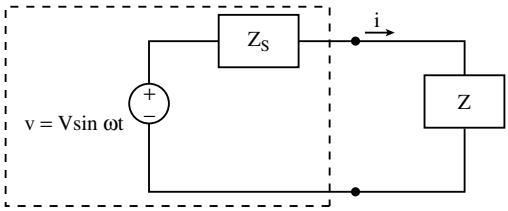


FIGURE 7.82 Z_s is fixed, and Z is to be chosen so that maximum average power is transferred to Z.

obtain the maximum average power transferred to the load [2]. Here, we assume that the resistance and reactance of the load can be independently adjusted. Let the source impedance be:

$$Z_S(j\omega) = R_S(\omega) + jX_S(\omega)$$

The load impedance is denoted as

$$Z(j\omega) = R(\omega) + jX(\omega)$$

Figure 7.82 depicts these impedances. We assume that all the elements, including the voltage source, within the box formed by the dotted lines are fixed. The voltage source is $v(t) = V\sin(\omega t)$, and thus $i(t) = I\sin(\omega t+\theta)$, where V and I are peak values and

$$\theta = -\tan^{-1}\left[\frac{X_S(\omega) + X(\omega)}{R_S(\omega) + R(\omega)}\right]$$

The average power delivered to Z is

$$P = I_{rms}^2 R(\omega)$$

where $I_{rms} = I/\sqrt{2}$ and

$$I = \frac{V}{\sqrt{\left[R_S(\omega) + R(\omega)\right]^2 + \left[X_S(\omega) + X(\omega)\right]^2}} \tag{7.121}$$

Thus, the average power delivered to Z can be written as

$$P = \frac{V_{rms}^2 R(\omega)}{\left[R_S(\omega) + R(\omega)\right]^2 + \left[X_S(\omega) + X(\omega)\right]^2} \tag{7.122}$$

To maximize P, we first note that the term $[X_s(\omega) + X(\omega)]^2$ is always positive, and so this term always contributes to a larger denominator unless it is zero. Thus, we set

$$X(\omega) = -X_S(\omega) \tag{7.123}$$

and (7.122) becomes

$$P = \frac{V_{rms}^2 R(\omega)}{\left[R_S(\omega) + R(\omega)\right]^2} \tag{7.124}$$

Second, we set the partial derivative with respect to $R(\omega)$ of the expression in (7.124) to zero to obtain

$$\frac{\partial P}{\partial R} = V_{rms}^2 \frac{\left(R_S + R\right)^2 - 2R\left(R_S + R\right)}{\left(R_S + R\right)^4} = 0 \tag{7.125}$$

Eq. (7.125) is satisfied for

$$R(\omega) = R_S(\omega) \tag{7.126}$$

and this value of $R(\omega)$ together with $X(\omega) = -X_s(\omega)$, yields maximum average power transferred to Z. Thus, we should adjust Z to:

$$Z(j\omega) = Z_S^*(j\omega) \tag{7.127}$$

and we obtain

$$P_{max} = \frac{V_{rms}^2}{4R(\omega)} \tag{7.128}$$

Example 28. Find Z for the network in Figure 7.83 so that maximum average power is transferred to Z. Determine the value of P_{max}.

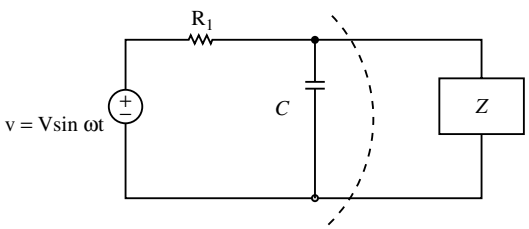


FIGURE 7.83 Circuit for Example 28.

Solution. We first obtain the Thévenin equivalent of the circuit to the left of the dotted arc in Figure 7.83 in order to reduce the circuit to the form of Figure 7.82.

$$\underline{V_{TH}} = \frac{\underline{V}}{1 + j\omega R_1 C}$$

$$Z_{TH} = \frac{R_1}{1 + j\omega R_1 C}$$

Thus,

$$Z = Z_{TH}^* = \frac{R_1}{1 - j\omega R_1 C} = \frac{R_1}{1 + \dfrac{\omega R_1 C}{j}}$$

$$= \frac{jR_1}{j + \omega R_1 C} = \frac{j\dfrac{1}{\omega C}R_1}{R_1 + j\dfrac{1}{\omega C}}$$

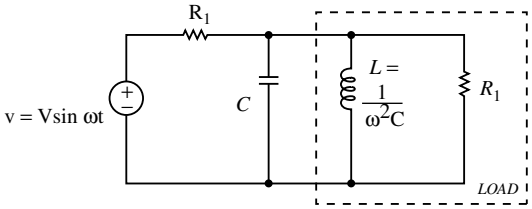


FIGURE 7.84 Circuit with load chosen to obtain maximum average power.

The term $j(\omega C)$ appears inductive (at a single frequency), and so we equate it to $j\omega L$ to obtain:

$$L = \frac{1}{\omega^2 C}$$

The impedance Z is therefore formed by the parallel connection of a resistor R_1 with the inductor L. Figure 7.84 depicts the resulting circuit. To determine P_{max}, we note that the capacitor and inductor constitute a parallel circuit which is resonant at the frequency of excitation. It therefore appears as an open circuit to the source. Thus, P_{max} is easily obtained as:

$$P_{\max} = I_{rms}^2 R_1 = \frac{V^2}{8R_1}$$

where V is the peak value of $v(t)$. □

Suppose Z is fixed and Z_s is adjustable in Figure 7.82. What should Z_s be so that maximum average power is delivered to Z? This is a problem that is applicable in the design of electronic amplifiers. The average power delivered to Z is given by (7.122), and to maximize P, we set $X_s(\omega) = -X(\omega)$ as before. We therefore obtain (7.124) again; but if R_s is adjustable instead of R, we see from (7.124) that P_{max} is obtained when R_s equals zero.

Acknowledgments

The author conveys his gratitude to Dr. Jacek Zurada, Mr. Tongfeng Qian, and to Dr. K. Wang for their help in proofreading this manuscript, and to Dr. Zbigniew J. Lata and Mr. Peichu (Peter) Sheng for producing the drawings.

References

[1] A. Budak, *Circuit Theory Fundamentals and Applications*, 2nd ed., Englewood Cliffs, NJ: Prentice Hall, 1987.
[2] L. P. Huelsman, *Basic Circuit Theory with Digital Computations*, Englewood Cliffs, NJ: Prentice Hall, 1972.
[3] L. P. Huelsman, *Basic Circuit Theory*, 3rd ed., Englewood Cliffs, NJ: Prentice Hall, 1991.
[4] S. Karni, *Applied Circuit Analysis*, New York: John Wiley & Sons, 1988.
[5] L. Weinberg, *Network Analysis and Synthesis*, New York: McGraw-Hill, 1962.

8

Symbolic Analysis

Benedykt S. Rodanski
University of Technology, Sydney

Marwan M. Hassoun
Iowa State University

8.1 Introduction and Definition

Symbolic circuit analysis, simply stated, is a term that describes the process of studying the behavior of electrical circuits using symbols instead of, or in conjunction with, numerical values. As an example to illustrate the concept, consider the input resistance of the simple circuit in Figure 8.1. Analyzing t**8**-he circuit using the unique symbols for each resistor without assigning any numerical values to them yields the input resistance of the circuit in the form:

$$\frac{V_{in}}{I_{in}} = \frac{R_1R_2 + R_1R_3 + R_1R_4 + R_2R_3 + R_2R_4}{R_2 + R_3 + R_4} \tag{8.1}$$

Equation (8.1) is the symbolic expression for the input resistance of the circuit in Figure 8.1.

The formal definition of symbolic circuit analysis can be written as:

Definition 1. Symbolic circuit analysis is the process of producing an expression that describes a certain behavioral aspect of the circuit with one, some, or all the circuit elements represented as symbols.

The idea of symbolic circuit analysis is not new; engineers and scientists have been using the process to study circuits since the inception of the concept of circuits. Every engineer has used symbolic circuit analysis during his or her education process. Most engineers still use it in their everyday job functions. As an example, all electrical engineers have symbolically analyzed the circuit in Figure 8.2. The equivalent resistance between nodes i and j is known to be:

$$\frac{1}{R_{ij}} = \frac{1}{R_1} + \frac{1}{R_2}$$

or

$$R_{ij} = \frac{R_1R_2}{R_1 + R_2}$$

This is the most primitive form of symbolic circuit analysis.

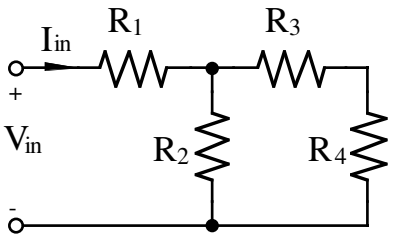


FIGURE 8.1 Symbolic circuit analysis example.

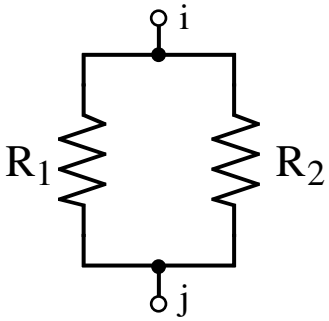


FIGURE 8.2 Common symbolic analysis problem.

The basic justification for performing symbolic analysis rather than numerical analysis on a circuit can be illustrated by considering the circuit in Figure 8.1 again. Assume that the values of all the resistances R_1 through R_4 are given as 1Ω and that the input resistance was analyzed numerically. The result obtained would be

$$\frac{V_{in}}{I_{in}} = \frac{5}{3} \approx 1.667\ \Omega \tag{8.2}$$

Now, consider the problem of increasing the input resistance of the circuit by adjusting only one of the resistor values. Equation (8.2) provides no insight into which resistor has the greatest impact on the input resistance. However, Eq. (8.1) clearly demonstrates that changing R_2, R_3, or R_4 would have very little impact on the input resistance because the terms appear in both the numerator and the denominator of the symbolic expression. It can also be observed that R_1 should be the resistor to change because it only appears in the numerator of the expression. Symbolic analysis has provided an insight into the problem.

From a circuit design perspective, numerical results from the simulation of a circuit can be obtained by evaluating the results of the symbolic analysis at a specific numerical point for each symbol. Ideally, only one simulation run is needed in order to analyze the circuit, and successive evaluations of the results replaces the need for any extra iterations through the simulator. Other applications include sensitivity analysis, circuit stability analysis, device modeling and circuit optimization [5, 18, 32, 41].

Although the previous "hand calculations" and somewhat trivial examples are used to illustrate symbolic circuit analysis, the thrust of the methods developed for symbolic analysis are aimed at computer implementations that are capable of symbolically analyzing circuits that cannot be analyzed "by hand." Several such implementations have been developed over the years [4, 10, 12, 15, 17, 22, 25, 26, 28, 29, 31, 34, 35, 37, 38, 47, 48, 53, 54, 56–61, 66].

Symbolic circuit analysis, referred to simply as symbolic analysis for the rest of this section, in its current form is limited to linear,[1] lumped, and time-invariant[2] networks. The scope of the analysis is

[1]Some references are made to the ability to analyze "weakly nonlinear" circuits [18, 63]; however, the actual symbolic analysis is performed on a linearized model of the weakly nonlinear circuit. Other techniques are applicable to circuits with only a single strongly nonlinear variable [65].

[2]One method is reported in Reference [36] that is briefly discussed in Section 8.6 that does deal with a limited class of time-variant networks.

primarily concentrated in the frequency domain, both *s*-domain [10, 12, 15, 17, 22, 25, 28, 29, 34, 38, 43, 47, 48, 54, 56, 59–61, 66] and *z*-domain [4, 31, 35, 44]; however, the predominant development has been in the *s*-domain. Also, recent work has expanded symbolic analysis into the time domain [3, 24, 36]. The next few subsections will discuss the basic methods used in symbolic analysis for mainly *s*-domain frequency analysis. However, Section 8.6 highlights the currently known time-domain techniques.

8.2 Frequency-Domain Analysis

Traditional symbolic circuit analysis is performed in the frequency domain where the results are in terms of the frequency variable *s*. The main goal of performing symbolic analysis on a circuit in the frequency domain is to obtain a symbolic transfer function of the form

$$H(s,\mathbf{x}) = \frac{N(s,\mathbf{x})}{D(s,\mathbf{x})}, \quad \mathbf{x} = \begin{bmatrix} x_1 & x_2 & \dots & x_p \end{bmatrix}, \quad p \le p_{all} \tag{8.3}$$

The expression is a rational function of the complex frequency variable s, and the variables x_1 through x_p representing the variable circuit elements, where p is the number of variable circuit elements and p_{all} is the total number of circuit elements. Both the numerator and the denominator of $H(s,\mathbf{x})$ are polynomials in s with real coefficients. Therefore, we can write

$$H(s,\mathbf{x}) = \frac{\sum_{i=0}^{m} a_i(\mathbf{x}) s^i}{\sum_{i=0}^{n} b_i(\mathbf{x}) s^i} = \frac{\prod_{i=1}^{m} \left[s - z_i(\mathbf{x}) \right]}{\prod_{i=1}^{n} \left[s - p_i(\mathbf{x}) \right]}$$

Most symbolic methods to date concentrate on the first form of $H(s,\mathbf{x})$ and several algorithms exist to obtain coefficients $a_i(\mathbf{x})$ and $b_i(\mathbf{x})$ in fully symbolic, partially symbolic (semi-symbolic), or numerical form. The zero/pole representation of $H(s,\mathbf{x})$, although more useful in gaining insight into circuit behavior, proved to be very difficult to obtain in symbolic form for anything but very simple circuits. For large circuits, various approximation techniques must be employed [9, 26].

A more recent approach to representing the above network function emerged in the 1980s and is based on a decomposed hierarchical form of Eq. (8.3) [22, 25, 51, 61, 62]. This hierarchical representation is referred to as a *sequence of expressions* representation to distinguish it from the *single expression* representation of Eq. (8.3) and is addressed in Section 8.4.

Several methodologies exist to perform symbolic analysis in the frequency domain. The early work was to produce a transfer function $H(s)$ with the frequency variable s being the only symbolic variable. Computer programs with these capabilities include: CORNAP [54] and NASAP [47]. The interest in symbolic analysis today is in the more general case when some or all of the circuit elements are represented by symbolic variables. The methods developed for this type of analysis fall under one of the following categories:

Traditional methods (single expression):

1. Tree enumeration methods
   - Single graph methods
   - Two graph methods
2. Signal flow graph methods
3. Parameter extraction methods
   - Modified nodal analysis-based methods
   - Tableau formulation-based methods
4. Interpolation method

Hierarchical methods (sequence of expressions):

1. Signal flow graph methods
2. Modified nodal analysis-based methods

The preceding classification includes the exact methods only. For large circuits, the traditional methods suffer from exponential growth of the number of terms in the formula with circuit size. If a certain degree of error is allowed, it may be possible to simplify the expression considerably, by including only the most significant terms. Several *approximate symbolic methods* have been investigated [26, 28, 69].

The next three sections discuss the basic theory for the above methods. Circuit examples are illustrated for all major methods except for the interpolation method due to its limited current usage[3] and its inability to analyze fully symbolic circuits.

8.3 Traditional Methods (Single Expressions)

This class of methods attempts to produce a single transfer function in the form of Eq. (8.3). The major advantage of having a symbolic expression in that form is the insight that can be gained by observing the terms in both the numerator and the denominator. The effects of the different terms can, perhaps, be determined by inspection. This process is valid for the cases where relatively few symbolic terms are in the expression.

Before indulging in the explanation of the different methods covered by this class, some definition of terms is in order.

Definition 2. *$RLCg_m$ circuit* is one that may contain only resistors, inductors, capacitors, and voltage-controlled current sources with the gain (transconductance) designated as g_m.

Definition 3. *Term cancellations* is the process in which two equal symbolic terms cancel out each other in the symbolic expression. This can happen in one of two ways: by having two equal terms with opposite signs added together, or by having two equal terms (regardless of their signs) divided by each other. For example, the equation

$$\frac{ab(ab+cd)-ab(cd-ef)}{ab(cd-gh)} \tag{8.4}$$

where a, b, c, d, e, f, g, and h are symbolic terms, can be reduced by observing that the terms ab in the numerator and denominator cancel each other and the terms $+cd$ and $-cd$ cancel each other in the numerator. The result is:

$$\frac{ab+ef}{cd-gh} \tag{8.5}$$

Definition 4. *Cancellation-free:* Equation (8.4) is said to be a cancellation-free equation (that is, no possible cancellations exist in the expression) while Eq. (8.5) is not.

Definition 5. *Cancellation-free algorithm*: The process of term cancellation can occur during the execution of an algorithm where a cancellation-free equation is generated directly instead of generating an expression with possible term cancellations in it. Cancellation-free algorithms are more desirable because, otherwise, an overhead is needed to generate and keep the terms that are to be canceled later.

[3]The main applications of the polynomial interpolation method in symbolic analysis are currently in numerical reference generation for symbolic approximation [14] and calculation of numerical coefficients in semi-symbolic analysis [50].

The different methods that fall under the traditional class are explained next.

1. The tree enumeration methods

Several programs have been produced based on this method [6, 16, 42, 46]. Practical implementations of the method can only handle small circuits in the range of 15 nodes and 30 branches [7]. The main reason is the exponential growth in the number of symbolic terms generated. The method can only handle one type of controlled source, namely, voltage controlled current sources. So only RLCg_m circuits can be analyzed. Also, the method does not produce any symbolic term cancellations for RLC circuits, and produces only a few for RLCg_m circuits.

The basic idea of the tree enumeration method is to construct an augmented circuit (a slightly modified version of the original circuit), its associated directed graph, and then enumerating all the directed trees of the graph. The admittance products of these trees are then used to find the node admittance matrix determinant and cofactors (the matrix itself is never constructed) to produce the required symbolic transfer functions. For a circuit with n nodes (with node n designated as the reference node) where the input is an excitation between nodes 1 and n and the output is taken between nodes 2 and n, the transfer functions of the circuit can be written as:

$$Z_{in} = \frac{V_1}{I_1} = \frac{\Delta_{11}}{\Delta} \tag{8.6}$$

$$\frac{V_o}{I_{in}} = \frac{V_2}{I_1} = \frac{\Delta_{12}}{\Delta} \tag{8.7}$$

$$\frac{V_o}{V_{in}} = \frac{V_2}{V_1} = \frac{\Delta_{12}}{\Delta_{11}} \tag{8.8}$$

where Δ is the determinant of the node admittance matrix $\mathbf{Y}_n$ (dimension n-1 × n-1) and Δ_{ij} is the ijth cofactor of $\mathbf{Y}_n$. It can be shown that a simple method for obtaining Δ, Δ_{11}, and Δ_{12} is to construct another circuit comprised of the original circuit with an extra admittance $\hat{y}_s$ in parallel with a voltage controlled current source, $\hat{g}_m V_2$, connected across the input terminals (nodes 1 and n). The determinant of $\hat{\mathbf{Y}}_n$ (the node admittance matrix for the new, slightly modified, circuit) can be written as:

$$\hat{\Delta} = \Delta + \hat{y}_S \Delta_{11} + \hat{g}_m \Delta_{12} \tag{8.9}$$

This simple trick allows the construction of the determinant expression of the original circuit and its two needed cofactors by simply formulating the expression for the new augmented circuit. Example 8.1 below illustrates this process.

The basic steps of the tree enumeration algorithm are (condensed from [7]):

1. Construct the augmented circuit from the original circuit by adding an admittance $\hat{y}_s$ and a transconductance $\hat{g}_m V_2$, in parallel between the input node and the reference node.
2. Construct a directed graph $\mathbf{G}_{\text{ind}}$ associated with the augmented circuit. The stamps used to generate $\mathbf{G}_{\text{ind}}$ are illustrated in Figure 8.3.
3. Find all directed trees for $\mathbf{G}_{\text{ind}}$. A directed tree rooted at node i is a subgraph of $\mathbf{G}_{\text{ind}}$ with node i having no incoming branches and each other node having exactly one incoming branch.
4. Find the admittance product for each directed tree. An admittance product of a directed tree is simply a term that is the product of all the weights of the branches in that tree.
5. Apply the following theorem:

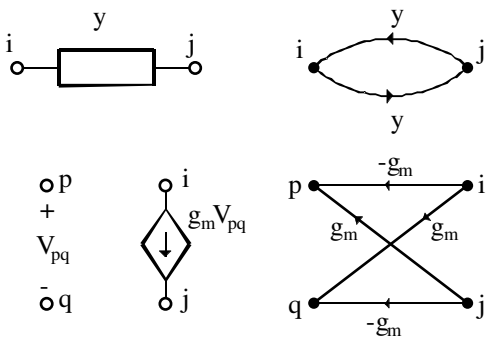


FIGURE 8.3 Element stamps for generating $\mathbf{G}_{ind}$.

Theorem 8.1 [7]: For any RLCg_m circuit, the determinant of the node admittance matrix (with any node as the reference node) is equal to the sum of all directed tree admittance products of $\mathbf{G}_{ind}$ (with any node as the root).

In other words

$$\hat{\Delta} = \sum \text{tree admittance products} \tag{8.10}$$

Arranging Eq. (8.10) in the form of Eq. (8.9) results in the necessary determinant and cofactors of the original circuit and the required transfer functions are generated from Eqs. (8.6), (8.7), and (8.8).

Example 1. A circuit and its augmented counterpart are illustrated in Figure 8.4. The circuit is the small-signal model of a simple inverting CMOS amplifier, shown with the coupling capacitance C_C taken into account. Figure 8.5 depicts the directed graph associated with the augmented circuit constructed using the rules in Figure 8.3. The figure also presents all the directed trees rooted at node 3 of the graph. Parallel branches heading in the same direction are combined into one branch with a weight equal to the sum of the weights of the individual parallel branches.

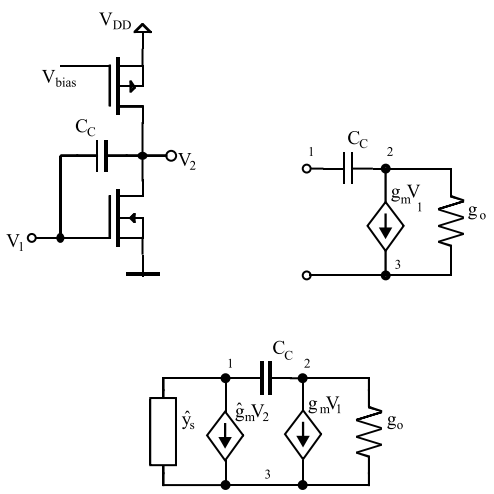


FIGURE 8.4 Circuit of Example 8.1 and its augmented equivalent diagram.

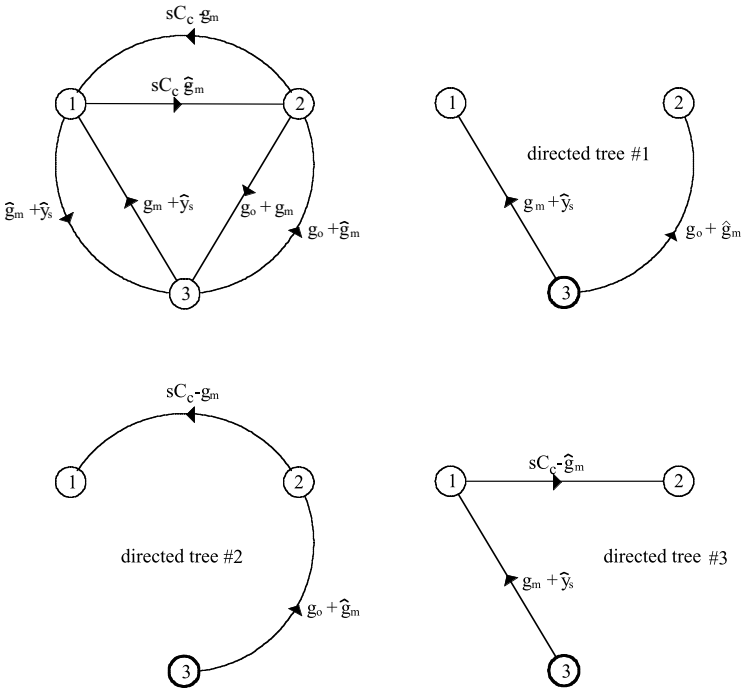


FIGURE 8.5 Graph and its directed trees of Example 8.1.

Applying Eq. (8.10) and rearranging the terms results in:

$$\begin{aligned}\hat{\Delta} &= \left(g_m + \hat{y}_S\right)\left(g_o + \hat{g}_m\right) + \left(sC_C - g_m\right)\left(g_o + \hat{g}_m\right) + \left(g_m + \hat{y}_S\right)\left(sC_C - \hat{g}_m\right) \\ &= \underbrace{sC_C\left(g_m + g_o\right)}_{\Delta} + \hat{y}_S\underbrace{\left(sC_C - g_o\right)}_{\Delta_{11}} + \hat{g}_m\underbrace{\left(sC_C - g_m\right)}_{\Delta_{12}}\end{aligned} \tag{8.11}$$

Note the fact that Eq. (8.11), which is the direct result of the algorithm, is not cancellation-free. Some terms cancel out to result in the determinant of the original circuit and its two cofactors of interest. The final transfer functions can be obtained readily by substituting the preceding results into Eq. (8.6) through (8.8).

2. The signal flow graph method

Two types of flow graphs are used in symbolic analysis. The first is referred to as a Mason's SFG and the second as Coates graph. Mason's SFG is by far a more popular and well-known SFG that has been used extensively in symbolic analysis among other controls applications. Both the Mason's SFG and the Coates graph are used as a basis for hierarchical symbolic analysis. However, the Coates graph was introduced to symbolic analysis by Starzyk and Konczykowska [61] solely for the purpose of performing hierarchical symbolic analysis. This section covers the Mason's SFG only.

The symbolic methods developed here are based on the idea formalized by Mason [45] in the 1950s. Formulation of the signal flowgraph and then the evaluation of the gain formula associated with it (Mason's formula) is the basis for symbolic analysis using this method. This method is used in the publicly available programs NASAP [47, 49] and SNAP [38]. The method has the same circuit size limitations as the tree enumeration method due to the exponential growth in the number of symbolic terms. However,

the signal flowgraph method allows all four types of controlled sources to be analyzed which made it a more popular method for symbolic analysis. The method is not cancellation-free, which contributes to the circuit size limitation mentioned earlier. An improved signal flowgraph method that avoids term cancellations was described in [48].

The analysis process of a circuit consists of two parts: the first is constructing the SFG for the given circuit and the second is to perform the analysis on the SFG. Some definitions are needed before proceeding to the details of these two parts.

Definition 6. *Signal Flow Graph:* An SFG is a weighted directed graph representing a system of simultaneous linear equations. Each node (x_i) in the SFG represents a circuit variable (node voltage, branch voltage, branch current, capacitor charge, or inductor flux) and each branch weight (w_{ij}) represents a coefficient relating x_i to x_j.

Every node in the SFG can be looked at as a summer. For a node x_k with m incoming branches

$$x_k = \sum_i w_{ik} x_i \tag{8.12}$$

where i spans the indices of all incoming branches from x_i to x_k.

Definition 7. *Path Weight:* The weight of a path from x_i to x_j (P_{ij}) is the product of all the branch weights in the path.

Definition 8. *Loop Weight:* The weight of a loop is the product of all the branch weights in that loop. This also holds for a loop with only one branch in it (self-loop).

Definition 9. *nth Order Loop:* An nth order loop is a set of n loops that have no common nodes between any two of them. The weight of an nth order loop is the product of the weights of all n loops.

Any transfer function x_j/x_i, where x_i is a source node, can be found by the application of Mason's formula:

$$\frac{x_j}{x_i} = \frac{1}{\Delta} \sum_k P_k \Delta_k \tag{8.13}$$

where

$$\begin{aligned} \Delta = 1 &- \sum_{\text{all}} \text{directed loop weights} \\ &+ \sum_{all} \text{2nd-order loop weights} \\ &- \sum_{all} \text{3rd-order loop weights} \\ &+ \cdots \end{aligned} \tag{8.14}$$

$$\begin{aligned} P_k &= \text{weight of the } k\text{th path from the source node } x_i \text{ to } x_j \\ \Delta_k &= \Delta \text{ with all loop contributions that are touching } P_k \text{ eliminated} \end{aligned} \tag{8.15}$$

The use of the preceding equations can be illustrated via an example.

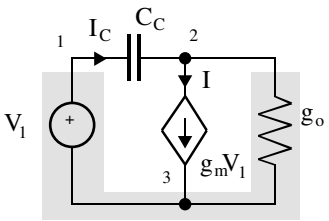


FIGURE 8.6 Circuit for Example 8.2 with its tree highlighted.

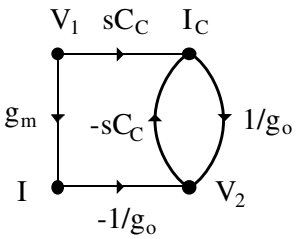


FIGURE 8.7 SFG for Example 8.2.

Example 2. Consider the circuit in Figure 8.6. The formulation of the SFG for this circuit takes on the following steps:

1. Find a tree and a co-tree of the circuit such that all current sources are in the co-tree and all voltage sources are in the tree.
2. Use Kirchhoff's current law (KCL), branch admittances, and tree branch voltages to find an expression for every co-tree link current. In the case of a controlled source, simply use the branch relationship. For the previous example, this yields:

$$I_C = sC_C\left(V_1 - V_2\right) = sC_CV_1 - sC_CV_2$$

$$I = g_mV_1$$

3. Use Kirchhoff's voltage law (KVL), branch impedances, and co-tree link currents to find an expression for every tree branch voltage. In the case of a controlled source, simply use the branch relationship. For the previous example, this yields:

$$V_{g_o} = V_2 = \frac{1}{g_o}\left(-I + I_C\right)$$

4. Create the SFG by drawing a node for each current source, voltage source, tree branch voltage, and co-tree link current.
5. Use Eq. (8.12) to draw the branches between the nodes that realize the linear equations developed in the previous steps.

Figure 8.7 is the result of executing the preceding steps on the example circuit. This formulation is referred to as the compact SFG. Any other variables that are linear combinations of the variables in the SFG (e.g., node voltages) can be added to the SFG by simply adding the extra node and implementing the linear relationship using SFG branches. A more detailed discussion of SFGs can be found in [7] and [40].

Now applying Eqs. (8.14) and (8.15) yields:

$$P_1 = -\frac{g_m}{g_o},\quad P_2 = \frac{sC_C}{g_o},\quad L_1 = -\frac{sC_C}{g_o},\quad \Delta = 1 - \left(-\frac{sC_C}{g_o}\right),\quad \Delta_1 = 1,\quad \Delta_2 = 1$$

Equation (8.13) then produces the final transfer function

$$\frac{V_2}{V_1} = \frac{1}{1+\dfrac{sC_C}{g_o}}\left(-\frac{g_m}{g_o}+\frac{sC_C}{g_o}\right) = \frac{sC_C - g_m}{sC_C + g_o}$$

3. The parameter extraction method

This method is best suited when few parameters in a circuit are symbolic while the rest of the parameters are in numeric form (*s* being one of the symbolic variables). The method was introduced in 1973 [2]. Other variations on the method were proposed later in [50, 56, 59]. The advantage of the method is that it is directly related to the basic determinant properties of widely used equation formulation methods such as the modified nodal method [27] and the tableau method [21]. As the name of the method implies, it provides a mechanism for extracting the symbolic parameters out of the matrix formulation, breaking the matrix solution problem into a numeric part and a symbolic part. The numeric part can then be solved using any number of standard techniques and recombined with the extracted symbolic part. The method has the advantage of being able to handle larger circuits than the previously discussed fully symbolic methods if only a few parameters are represented symbolically. If the number of symbolic parameters in a circuit is high, the method will exhibit the same exponential growth in the number of symbolic terms generated and will have the same circuit size limitations as the other algorithms previously discussed.

The method does not limit the type of matrix formulation used to analyze the circuit. However, the extraction rules depend on the pattern of the symbolic parameters in the matrix. Alderson and Lin [1] use the indefinite admittance matrix as the basis of the analysis and the rules depend on the appearance of a symbolic parameter in four locations in the matrix: (*i*,*i*), (*i*,*j*), (*j*,*i*), and (*j*,*j*). Singhal and Vlach [59] use the tableau equations and can handle a symbolic parameter that only appears once in the matrix. Sannuti and Puri [56] force the symbolic parameters to appear only on the diagonal using a two-graph method [7] to write the tableau equations. The parameter extraction method was further simplified in [50], where the formula is given to calculate a coefficient (generally a polynomial in *s*) at every symbol combination. Some invalid symbol combinations (i.e., the ones that do not appear in the final formula) can be eliminated before calculations by topological considerations. To illustrate both approaches to parameter extraction, this section presents the indefinite admittance matrix (IAM) formulation and the most recent two-graph method. Details of other formulations can be found in [40, 48, 56, 59].

Indefinite Admittance Matrix Approach

One of the basic properties of the IAM is the symmetric nature of the entries sometimes referred to as *quadrantal* entries [7, 40]. A symbolic variable α will always appear in four places in the indefinite admittance matrix, $+\alpha$ in entries (i, k) and (j, m), and $-\alpha$ in entries (i, m) and (j, k) as demonstrated in the following equation:

$$\begin{array}{c} \\ i \\ \\ j \\ \\ \end{array}\begin{array}{c} \begin{array}{ccc} k & & m \end{array} \\ \begin{bmatrix} \vdots & & \vdots \\ \alpha & \cdots & -\alpha \\ \vdots & & \vdots \\ -\alpha & \cdots & -\alpha \\ \vdots & & \vdots \end{bmatrix} \end{array}$$

where $i \neq j$ and $k \neq m$. For the case of an admittance y between nodes i and j, we have $k = i$ and $j = m$. The basic process of extracting the parameter (the symbol) α can be performed by applying the following equation [2, 7]:

$$\text{cofactor of } \mathbf{Y}_{ind} = \text{cofactor of } \mathbf{Y}_{ind,\alpha=0} + (-1)^{j+m}\alpha\left(\text{cofactor of } \mathbf{Y}_{\alpha}\right) \tag{8.16}$$

where $\mathbf{Y}_{\alpha}$ is a matrix that does not contain α and is obtained by:

1. Adding row j to row i
2. Adding column m to column k
3. Deleting row j and column m

For the case where several symbols exist, the previous extraction process can be repeated and would result in

$$\text{cof}\left(\mathbf{Y}_{ind}\right) = \sum_{j} P_j \,\text{cof}\left(\mathbf{Y}_j\right)$$

where P_j is some product of symbolic parameters including the sign and $\mathbf{Y}_j$ is a matrix with the frequency variable s, possibly being the only symbolic variable. The cofactor of $\mathbf{Y}_j$ may be evaluated using any of the usual evaluation methods [7, 64]. Programs implementing this technique include NAPPE2 [40] and SAPWIN [37].

Example 4 [7]. Consider the resistive circuit in Figure 8.8. The goal is to find the input impedance Z_{14} using the parameter extraction method where g_m is the only symbolic variable in the circuit. In order to use Eqs. (8.6) and (8.9), an admittance $\hat{y}_s$ is added across the input terminals of the circuit to create the augmented circuit.

The IAM is then written as (conductances in siemens [S])

$$\hat{\mathbf{Y}}_{ind} = \begin{bmatrix} 6+\hat{y}_S & -5 & -1 & -\hat{y}_S \\ g_m - 5 & 15.1 & -g_m - 10 & -0.1 \\ -g_m - 1 & -10 & g_m + 13 & -2 \\ -\hat{y}_S & -0.1 & -2 & \hat{y}_S + 2.1 \end{bmatrix}$$

Applying Eq. (8.16) to extract $\hat{y}_s$ results in

$$\text{cof}\left(\hat{\mathbf{Y}}_{ind}\right) = \text{cof}\begin{bmatrix} 6 & -5 & -1 & 0 \\ g_m - 5 & 15.1 & -g_m - 10 & -0.1 \\ -g_m - 1 & -10 & g_m + 13 & -2 \\ 0 & -0.1 & -2 & 2.1 \end{bmatrix} + \hat{y}_S \,\text{cof}\begin{bmatrix} 8.1 & -5.1 & -3 \\ g_m - 5.1 & 15.1 & -g_m - 10 \\ -g_m - 3 & -10 & g_m + 13 \end{bmatrix}$$

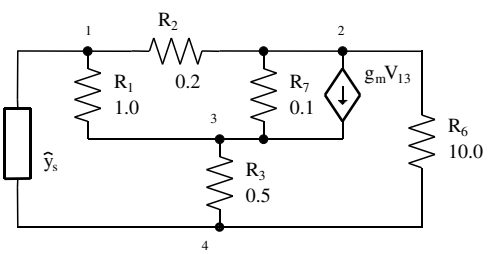


FIGURE 8.8 Circuit for the parameter extraction method (resistances in ohms [Ω]).

Applying Eq. (8.16) again to extract g_m yields

$$\operatorname{cof}\left(\hat{\mathbf{Y}}_{ind}\right) = \operatorname{cof}\begin{bmatrix} 6 & -5 & -1 & 0 \\ -5 & 15.1 & -10 & -0.1 \\ -1 & -10 & +13 & -2 \\ 0 & -0.1 & -2 & 2.1 \end{bmatrix} + g_m \operatorname{cof}\begin{bmatrix} 5 & -5 & 0 \\ -3 & 5.1 & -2.1 \\ -2 & -0.1 & 2.1 \end{bmatrix}$$

$$+\hat{y}_S \operatorname{cof}\begin{bmatrix} 8.1 & -5.1 & -3 \\ -5.1 & 15.1 & -10 \\ -3 & -10 & 13 \end{bmatrix} + \hat{y}_S g_m \operatorname{cof}\begin{bmatrix} 5.1 & -5.1 \\ -5.1 & 5.1 \end{bmatrix}$$

After evaluating the cofactors numerically, the equation reduces to

$$\operatorname{cof}\left(\hat{\mathbf{Y}}_{ind}\right) = 137.7 + 10.5g_m + 96.3\hat{y}_S + 5.1\hat{y}_S g_m$$

From Eq. (8.9), this results in

$$Z_{14} = \frac{\Delta_{11}}{\Delta} = \frac{96.3 + 5.1g_m}{137.7 + 10.5g_m}$$

Two-Graph-Based Tableau Approach [50]

This approach also employs the circuit augmentation by $\hat{y}_s$ and $\hat{g}_m V_o$, as in the tree enumeration method. It calls for the construction of two graphs: the voltage graph (G_V or V-graph) and the current graph (G_I or I-graph). For the purpose of parameter extraction (as well as generation of approximate symbolic expressions; see Section 8.5), it is required that both graphs have the same number of nodes (n). This means that the method can be directly applied only to RLCg_m circuits. (All basic circuit components, including ideal op amps, can be handled by this approach after some circuit transformations [40, 64]. For the sake of simplicity, however, only RLCg_m circuits will be considered in this presentation.) The two graphs are constructed based on the element stamps shown in Figure 8.9. Once the two graphs are constructed, a *common spanning tree* (i.e., a set of n-1 branches that form a spanning tree in both voltage and current graphs) is chosen. Choosing the common spanning tree (referred to just as "tree" in the remainder of this section) uniquely determines the co-tree in each graph.

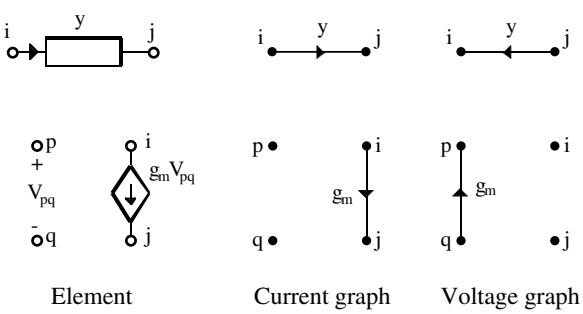


FIGURE 8.9 Element stamps for generating G_V and G_I.

The tableau equation for such a network can be written as

$$\mathbf{Hx} = \begin{bmatrix} \mathbf{1} & \mathbf{0} & -\mathbf{Z}_T & \mathbf{0} \\ \mathbf{B}_T & \mathbf{1} & \mathbf{0} & \mathbf{0} \\ \mathbf{0} & \mathbf{0} & \mathbf{1} & \mathbf{Q}_C \\ \mathbf{0} & -\mathbf{Y}_C & \mathbf{0} & \mathbf{1} \end{bmatrix} \begin{bmatrix} \mathbf{V}_T \\ \mathbf{V}_C \\ \mathbf{I}_T \\ \mathbf{I}_C \end{bmatrix} = \mathbf{0} \tag{8.17}$$

The first and last row of the system matrix **H** in Eq. (8.17) consists of tree ($\bullet_T$) and co-tree ($\bullet_C$) branch voltage-current relationships, and the second and third rows consist of fundamental loop and fundamental cut-set equations for G_V and G_I, respectively.

Let the circuit have n nodes and b branches and contain k symbolic components (Y_{S1}, …, Y_{Sk}) in the co-tree branches (links) and l symbolic components (Z_{S1}, …, Z_{Sl}) in the tree branches; we define $w = b - n - k + 1$, $t = n - l - 1$. Diagonal matrices $\mathbf{Y}_C$ and $\mathbf{Z}_T$ can be partitioned as follows

$$\mathbf{Y}_C = \begin{bmatrix} \mathbf{Y}_{Cs} & \mathbf{0} \\ \mathbf{0} & \mathbf{Y}_{Cn} \end{bmatrix}, \quad \mathbf{Z}_T = \begin{bmatrix} \mathbf{Z}_{Ts} & \mathbf{0} \\ \mathbf{0} & \mathbf{Z}_{Tn} \end{bmatrix} \tag{8.18}$$

where subscript s denotes immitances of symbolic components and subscript n denotes immitances of components given numerically.

Matrices $\mathbf{B}_T$ (fundamental loop matrix in G_V) and $\mathbf{Q}_C$ (fundamental cut-set matrix in G_I) can also be partitioned as follows:

$$\mathbf{B}_T = \begin{bmatrix} \mathbf{B}_{11} & \mathbf{B}_{12} \\ \mathbf{B}_{21} & \mathbf{B}_{22} \end{bmatrix}, \quad \mathbf{Q}_C = \begin{bmatrix} \mathbf{Q}_{11} & \mathbf{Q}_{12} \\ \mathbf{Q}_{21} & \mathbf{Q}_{22} \end{bmatrix} \tag{8.19}$$

Rows of $\mathbf{B}_{11}$ and $\mathbf{B}_{12}$ correspond to symbolic co-tree branches (in G_V) and their columns correspond to symbolic and numeric tree branches, respectively. Rows of $\mathbf{B}_{21}$ and $\mathbf{B}_{22}$ correspond to numeric co-tree branches. Rows of $\mathbf{Q}_{11}$ and $\mathbf{Q}_{12}$ correspond to symbolic tree branches and their columns correspond to symbolic and numeric co-tree branches, respectively. Rows of $\mathbf{Q}_{21}$ and $\mathbf{Q}_{22}$ correspond to numeric tree branches. The submatrices are therefore of the following order: $\mathbf{B}_{11}$: $k \times l$, $\mathbf{B}_{22}$: $w \times t$, $\mathbf{Q}_{11}$: $l \times k$, $\mathbf{Q}_{22}$: $t \times w$.

Let $S_x = \{1, 2, \ldots, x\}$. For a given matrix **F** of order $a \times b$ let $\mathbf{F}(I_u, J_v)$ be the submatrix of **F** consisting of the rows and columns corresponding to the integers in the sets I_u, J_v, respectively. The sets $I_u = \{i_1, i_2, \ldots, i_u\}$ and $J_v = \{j_1, j_2, \ldots, j_v\}$ are subsets of S_a and S_b, respectively. Let us also introduce the following notation:

$$\mathbf{1}_d^c = \mathrm{diag}\begin{bmatrix} e_1 & e_2 & \cdots & e_d \end{bmatrix}; \quad c < d$$

$$e_x = \begin{cases} 0 \text{ for } x \in \{1, 2, \ldots, c\} \\ 1 \text{ for } x \in \{c+1, c+2, \ldots, d\} \end{cases}$$

The determinant of the system matrix **H** in Eq. (8.17), when some parameters take fixed numerical values, is

$$\begin{aligned} \det H = a &+ \sum_{J_v} b(\alpha_v) Z_{s_{j1}} Z_{s_{j2}} \cdots Z_{s_{jv}} + \sum_{I_u} c(\beta_u) Y_{s_{i1}} Y_{s_{i2}} \cdots Y_{s_{iu}} \\ &+ \sum_{I_u} \sum_{J_v} d(\alpha_v \beta_u) Z_{s_{j1}} Z_{s_{j2}} \cdots Z_{s_{jv}} Y_{s_{i1}} Y_{s_{i2}} \cdots Y_{s_{iu}} \end{aligned} \tag{8.20}$$

where the summations are taken over all possible symbol combinations α_v (symbolic tree elements) and β_u (symbolic co-tree elements), and the numerical coefficients are given by:

$$
\begin{aligned}
a &= \det\left[\mathbf{1}_w + \mathbf{B}'_{22}\left(-\mathbf{Q}'_{22}\right)\right] = \det\left[\mathbf{1}_t + \left(-\mathbf{Q}'_{22}\right)\mathbf{B}'_{22}\right] \\
b\left(\alpha_v\right) &= \det\left(\mathbf{1}^{v}_{t+v} + \begin{bmatrix} -\mathbf{Q}_{12}\left(J_v, I_w\right) \\ -\mathbf{Q}'_{22} \end{bmatrix}\begin{bmatrix} \mathbf{B}'_{21}\left(I_w, J_v\right) & \mathbf{B}'_{22} \end{bmatrix}\right) \\
c\left(\beta_u\right) &= \det\left(\mathbf{1}^{u}_{w+u} + \begin{bmatrix} \mathbf{B}_{12}\left(I_u, J_t\right) \\ \mathbf{B}'_{22} \end{bmatrix}\begin{bmatrix} -\mathbf{Q}'_{21}\left(J_t, I_u\right) & -\mathbf{Q}'_{22} \end{bmatrix}\right) \\
d\left(\alpha_v\beta_u\right) &= \det\left(\mathbf{1}^{v}_{v+t+u} + \begin{bmatrix} -\mathbf{Q}_{11}\left(J_v, I_u\right) & -\mathbf{Q}_{12}\left(J_v, I_w\right) \\ -\mathbf{Q}'_{21}\left(J_t, I_u\right) & -\mathbf{Q}'_{22} \\ \mathbf{1}_u & \mathbf{0} \end{bmatrix}\begin{bmatrix} \mathbf{B}_{11}\left(I_u, J_v\right) & \mathbf{B}_{12}\left(I_u, J_t\right) & -\mathbf{1}_u \\ \mathbf{B}'_{21}\left(I_w, J_v\right) & \mathbf{B}'_{22} & \mathbf{0} \end{bmatrix}\right)
\end{aligned}
\tag{8.21}
$$

In the preceding equations, **0** represents a zero matrix of appropriate order, and the submatrices $\mathbf{B}'_{ij}$ and $\mathbf{Q}'_{ij}$ are defined as:

$$
\begin{aligned}
\mathbf{B}'_{21}\left(I_w, J_v\right) &= \mathbf{Y}_{Cn}\mathbf{B}_{21}\left(I_w, J_v\right), \quad \mathbf{B}'_{22} = \mathbf{Y}_{Cn}\mathbf{B}_{22} \\
\mathbf{Q}'_{21}\left(J_t, I_u\right) &= \mathbf{Z}_{Tn}\mathbf{Q}_{21}\left(J_t, I_u\right), \quad \mathbf{Q}'_{22} = \mathbf{Z}_{Tn}\mathbf{Q}_{22}
\end{aligned}
\tag{8.22}
$$

where the submatrix $\mathbf{B}_{21}(I_w, J_v)$ is obtained from the submatrix $\mathbf{B}_{21}$ by including all of its rows and only columns corresponding to a particular combination (α_v) of symbolic tree elements; submatrix $\mathbf{Q}_{21}(J_t, I_u)$ is obtained from the submatrix $\mathbf{Q}_{21}$ by including all of its rows and only columns corresponding to a particular combination (β_u) of symbolic co-tree elements.

Application of Eqs. (8.20) and (8.21) for a circuit with m symbolic parameters requires, theoretically, the calculation of 2^m determinants. Not all of these determinants may need to be calculated due to the following property of the determinants in Eq. (8.21). If a set of symbolic tree elements (α_v) forms a cut-set in G_I (*symbolic tree cut-set*), then the corresponding coefficients $b(\alpha_v)$ and $d(\alpha_v\beta_u)$ in Eq. (8.20) equal to zero. Likewise, if the set of symbolic co-tree elements (β_u) forms a loop in G_V (*symbolic co-tree loop*), the corresponding coefficients $c(\beta_u)$ and $d(\alpha_v\beta_u)$ in Eq. (8.20) equal to zero.

Once the determinant det(**H**) is obtained from Eq. (8.20), the sorting scheme, identical to that expressed in Eq. (8.9), is applied and the required network function(s) can be calculated using Eqs. (8.6) through (8.8).

The main feature of this approach is the fact that each coefficient at a valid symbol combination is obtained directly by calculating a single, easily formulated determinant (a polynomial in s, in general case). The method was implemented in a computer program called UTSSNAP [52]. The following example illustrates this technique of parameter extraction.

Example 5. Consider again the circuit in Figure 8.8. Assume this time that two components, R_1 and g_m, are given symbolically. The goal is again to find the input impedance Z_{41} in a semi-symbolic form using the parameter extraction method based on the two-graph tableau formulation.

The voltage and current graphs of the circuit are shown in Figure 8.10. The common spanning tree chosen is $T = \{R_1, R_2, R_3\}$ with one symbolic element. For this circuit, we have: $n = 4$, $b = 7$, $k = 2$, $l = 1$, $w = 2$, and $t = 2$.

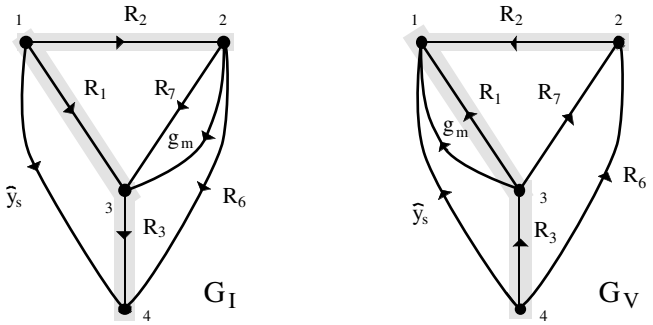


FIGURE 8.10 The current and voltage graphs for the circuit in Figure 8.8 with the common spanning tree highlighted.

The matrices $\mathbf{Y}_C$, $\mathbf{Z}_T$, $\mathbf{Q}_C$, and $\mathbf{B}_T$ can now be determined as:

$$\mathbf{Y}_C = \left[\begin{array}{cc|c} \hat{y}_S & & \\ & g_m & \\ \hline & & 0.1 \\ & & & 10 \end{array}\right] \qquad \mathbf{Z}_T = \left[\begin{array}{cc|c} R_1 & & \\ \hline & & 0.2 \\ & & & 0.5 \end{array}\right]$$

$$\mathbf{Q}_C = \left[\begin{array}{cc|cc} 1 & 1 & 1 & 1 \\ \hline 0 & -1 & -1 & -1 \\ 1 & 0 & 1 & 0 \end{array}\right] \qquad \mathbf{B}_T = \left[\begin{array}{c|cc} -1 & 0 & -1 \\ -1 & 0 & 0 \\ \hline -1 & 1 & -1 \\ -1 & 1 & 0 \end{array}\right]$$

Using Eq. (8.22), we can calculate matrices $\mathbf{B}'_{22}$ and $\mathbf{Q}'_{22}$:

$$\mathbf{B}'_{22} = \begin{bmatrix} 0.1 & 0 \\ 0 & 10 \end{bmatrix}\begin{bmatrix} 1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} 0.1 & -0.1 \\ 10 & 0 \end{bmatrix}, \quad \mathbf{Q}'_{22} = \begin{bmatrix} 0.2 & 0 \\ 0 & 0.5 \end{bmatrix}\begin{bmatrix} -1 & -1 \\ 1 & 0 \end{bmatrix} = \begin{bmatrix} -0.2 & -0.2 \\ 0.5 & 0 \end{bmatrix}$$

Now, applying Eq. (8.21), the coefficient a in Eq. (8.20) is calculated as:

$$a = \det\left(\begin{bmatrix} 1 & 0 \\ 0 & 1 \end{bmatrix} + \begin{bmatrix} 0.1 & -0.1 \\ 10 & 0 \end{bmatrix}\begin{bmatrix} 0.2 & 0.2 \\ -0.5 & 0 \end{bmatrix}\right) = \det\begin{bmatrix} 1.07 & 0.02 \\ 2 & 3 \end{bmatrix} = 3.17$$

Because only one symbolic tree element exists, namely R_1, we have: $\alpha_v = \{R_1\}$ and the associated sets: $J_v = \{1\}$, $I_w = \{1,2\}$. Using Eq. (8.22), we calculate

$$\mathbf{B}'_{21}(I_w, J_v) = \mathbf{Y}_{Cn}\mathbf{B}_{21}(I_w, J_v) = \begin{bmatrix} 0.1 & 0 \\ 0 & 10 \end{bmatrix}\begin{bmatrix} -1 \\ -1 \end{bmatrix} = \begin{bmatrix} -0.1 \\ -10 \end{bmatrix}$$

The coefficient $b(R_1)$ can now be obtained from:

$$b(R_1) = \det\left(\begin{bmatrix} 0 & 0 & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{bmatrix} + \left[\begin{array}{cc} -1 & -1 \\ \hline 0.2 & 0.2 \\ -0.5 & 0 \end{array}\right]\left[\begin{array}{c|cc} -0.1 & 0.1 & -0.1 \\ -10 & 10 & 0 \end{array}\right]\right)$$

$$= \det\begin{bmatrix} 10.1 & -10.1 & 0.1 \\ -2.02 & 3.02 & -0.02 \\ 0.05 & -0.05 & 1.05 \end{bmatrix} = 10.6$$

Other numerical coefficients in Eq. (8.20) are calculated in a similar way:

$$c(\hat{y}_S) = 1.51,\ c(g_m) = 0,\ c(\hat{y}_S g_m) = 0,\ d(R_1\hat{y}_S) = 8.12,\ d(R_1 g_m) = 1.05,\ d(R_1\hat{y}_S g_m) = 0.51$$

Adding all terms, sorting according to Eq. (8.9), and applying Eq. (8.6) finally results in:

$$Z_{41} = \frac{1.51 + 8.12R_1 + 0.51R_1 g_m}{3.17 + 10.6R_1 + 1.05R_1 g_m}$$

Matrices in Eq. (8.21) may contain terms dependent on the complex frequency s. Determinants of such matrices are polynomials in s as long as all matrix elements are of the form: $a = \alpha + s\beta$. An interpolation method may be used to calculate the coefficients of those polynomials. One such method is briefly described in the next paragraph.

4. The interpolation method

This method is best suited when s is the only symbolic variable. In such case, a transfer function has the rational form

$$H(s) = \frac{N(s)}{D(s)} = \frac{\sum_{i=0}^{m} a_i s^i}{\sum_{i=0}^{n} b_i s^i}$$

where $N(s)$ and $D(s)$ are polynomials in s with real coefficients and $m \le n$.

Coefficients of an nth-order polynomial

$$P(s) = \sum_{k=0}^{n} p_k s^k$$

can be obtained by calculating the value of $P(s)$ at $n + 1$ distinct points s_i and then solving the following set of equations:

$$\begin{bmatrix} 1 & s_0 & s_0^2 & \cdots & s_0^n \\ 1 & s_1 & s_1^2 & \cdots & s_1^n \\ & \vdots & & & \\ 1 & s_n & s_n^2 & \cdots & s_n^n \end{bmatrix}\begin{bmatrix} p_0 \\ p_1 \\ \vdots \\ p_n \end{bmatrix} = \begin{bmatrix} P(s_0) \\ P(s_1) \\ \vdots \\ P(s_n) \end{bmatrix} \tag{8.23}$$

Because the matrix in Eq. (8.23) is nonsingular, the unique solution exists. It is well known [58, 64] that for numerical accuracy and stability, the best choice of the interpolation points is a set of $q \geq n + 1$ points s_i uniformly spaced on the unit circle in the complex plane. Once all the values of $P(s_i)$ are known, the polynomial coefficients can be calculated through the discrete Fourier transform (DFT).

To apply this technique to the problem of finding a transfer function, let us assume that a circuit behavior is described by a linear equation

$$\mathbf{Ax} = \mathbf{b} \tag{8.24}$$

in which the coefficient matrix has entries of the form: $a = \alpha + s\beta$ (both the modified nodal and the tableau methods have this property). Then, each transfer function of such circuit has the same denominator $D(s) = |\mathbf{A}|$. If the circuit Eq. (8.24) is solved by LU factorization at $s = s_i$, both the transfer function $H(s_i)$ and its denominator $D(s_i)$ are obtained simultaneously. The value of the numerator is then calculated simply as $N(s_i) = H(s_i)D(s_i)$. Repeating this process for all points s_i ($i = 0, 1, \ldots, q$) and then applying the DFT to both sets of values, $D(s_i)$ and $N(s_i)$, gives the required coefficients of the numerator and denominator polynomials.

If the number of interpolation points is an integer power of 2 ($q = 2^k$), the method has the advantage that the fast Fourier transform can be used to find the coefficients. This greatly enhances the execution time [40]. The method has been extended to handle several symbolic variables in addition to s [58]. The program implementation [64] allows a maximum of five symbolic parameters in a circuit.

With the emergence of approximate symbolic analysis, the polynomial interpolation method has attracted new interest. (It is desirable to know the accurate numerical value of polynomial coefficients before one attempts an approximation.) Recently, a new adaptive scaling mechanism was proposed [14] that significantly increases the circuit size that can be handled accurately and efficiently.

Other classifications of symbolic methods have been reported [18]. These methods can be considered as variations on the previous basic four methods. The reported methods include elimination algorithms, recursive determinant-expansion algorithms, and nonrecursive nested-minors method. All three are based on the use of Cramer's rule to find the determinant and the cofactors of a matrix. Another reported class of algorithms uses Modified Nodal Analysis [27] as the basis of the analysis, sometimes referred to as a direct network approach [22, 36]. This class of methods is covered in the next section.

The first generation of computer programs available for symbolic circuit simulation based on these methods includes NASAP [47] and SNAP [38]. Research in the late 1980s and early 1990s produced newer symbolic analysis programs. These programs include ISSAC [18], SCAPP [22], ASAP [12], EASY [60], SYNAP [57], SAPEC [43], SAPWIN [37], SCYMBAL [31], GASCAP [29], SSPICE [66], and STAINS [53].

8.4 Hierarchical Methods (Sequence of Expressions)

All the methods presented in the previous section have circuit size limitations. The main problem is the exponential growth of the number of symbolic terms involved in the expression for the transfer function in Eq. (8.3) as the circuit gets larger. The solution to analyzing large-scale circuits lies in a total departure from the traditional procedure of trying to state the transfer function as a single expression and using a *sequence of expressions* (SoE) procedure instead. The idea is to produce a succession of small expressions with a backward hierarchical dependency on each other. The growth of the number of expressions in this case will be, at worst case, quadratic [22].

The advantage of having the transfer function stated in a single expression lies in the ability to gain insight to the relationship between the transfer function and the network elements by inspection [39]. For large expressions, though, this is not possible and the single expression loses that advantage. ISSAC [67], ASAP [13], SYNAP [57], and Analog Insydes [26] attempt to handle larger circuits by maintaining the single expression method and using circuit dependent approximation techniques. The tradeoff is

accuracy for insight. Therefore, the SoE approach is more suitable for accurately handling large-scale circuits. The following example illustrates the features of the sequence of expressions.

Example 6. Consider the resistance ladder network in Figure 8.11. The goal is to obtain the input impedance function of the network, $Z_{in} = V_{in}/I_{in}$. The single expression transfer function Z_4 is:

$$Z_4 = \frac{R_1R_3 + R_1R_4 + R_2R_3 + R_2R_4 + R_3R_4}{R_1 + R_2 + R_3}$$

The number of terms in the numerator and denominator are given by the Fibonacci numbers satisfying the following difference equation:

$$y_{k+2} = y_{k+1} + y_k; \quad k = 0,1,2,\ldots; \quad y_0 = 0,\ y_1 = 1$$

An explicit solution to the preceding equation is:

$$y_n = \frac{1}{\sqrt{5}}\left(\frac{1+\sqrt{5}}{2}\right)^n - \left(\frac{1-\sqrt{5}}{2}\right)^n \approx 0.168 \cdot 1.618^n \text{ for large } n$$

The solution demonstrates that the number of terms in Z_n increases exponentially with n. Any single expression transfer function has this inherent limitation.

Now, using the SoE procedure, the input impedance can be obtained from the following expressions:

$$Z_1 = R_1; \quad Z_2 = Z_1 + R_2; \quad Z_3 = \frac{Z_2R_3}{Z_2 + R_3}; \quad Z_4 = Z_3 + R_4$$

It is obvious for each additional resistance added, the sequence of expressions will grow by one expression, either of the form $Z_{i-1} + R_i$ or $Z_{i-1}R_i/Z_{i-1} + R_i$. The number of terms in the sequence of expressions can be calculated from the formula:

$$y_n = \begin{cases} 2.5n - 2 & \text{for } n \text{ even} \\ 2.5n - 1.5 & \text{for } n \text{ odd} \end{cases}$$

which exhibits a linear growth with respect to n. Therefore, to find the input impedance of a 100-resistor ladder network, the single expression methods would produce 7.9×10^{20} terms, which requires unrealistically huge computer storage capabilities. On the other hand, the SoE method would produce only 248 terms, which is even within the scope of some desk calculators.

Another advantage of the SoE is the number of arithmetic operations needed to evaluate the transfer function. To evaluate Z_9, for example, the single expression methods would require 302 multiplications

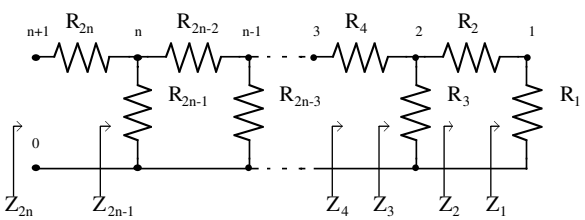


FIGURE 8.11 Resistive ladder network.

and 87 additions. The SoE method would only require eight multiplications and eight additions, a large reduction in computer evaluation time. All this makes the concept of symbolic circuit simulation of large-scale networks very possible.

Two topological analysis methods for symbolic simulation of large-scale circuits have been proposed in [61] and in [25]. The first method utilizes the SoE idea to obtain the transfer functions. The method operates on the Coates graph [8] representing the circuit. A partitioning is proposed onto the flowgraph and not the physical network. The second method also utilizes the sequence of expressions and a Mason's signal flow graph [45] representation of the circuit. The method makes use of partitioning on the physical level instead of on the graph level. Therefore, for a hierarchical circuit, the method can operate on the subcircuits in a hierarchical fashion in order to produce a final solution. The fundamentals of both signal flow graph methods were described in the previous section.

Another hierarchical approach is one that is based on Modified Nodal Analysis [27]. This method [22] exhibits a linear growth (for practical circuits) in the number of terms in the symbolic solutions. The analysis methodology introduces the concept of the RMNA (Reduced Modified Nodal Analysis) matrix. This allows the characterization of symbolic circuits in terms of only a small subset of the network variables (external variables) instead of the complete set of variables. The method was made even more effective by introducing a locally optimal pivot selection scheme during the reduction process [53]. For a circuit containing several identical[4] subcircuits, the analysis algorithm is most efficient when network partitioning is used. For other circuits, the best results (the most compact SoE) are obtained when the entire circuit is analyzed without partitioning.

The SoE generation process starts with the formulation of a symbolic Modified Node Admittance Matrix (MNAM) for a circuit [40, 64]. Then all internal variables are suppressed one by one using Gaussian elimination with locally optimal pivot selection. Each elimination step produces a series of expressions and modifies some entries in the remaining portion of the MNAM. When all internal variables are suppressed, the resulting matrix is known as the Reduced Modified Node Admittance Matrix (RMNAM). Usually it will be a 2×2 matrix of a two-port.[5] Most transfer functions of interest to a circuit designer can be represented by formulas involving the elements of RMNAM and the terminating admittances. A detailed discussion of the method can be found in [53]. Based on this approach, a computer program called STAINS was developed.

For a circuit with several identical subcircuits, the reduction process is first applied to all internal variables[6] of the subcircuit, resulting in an intermediate RMNAM describing the subcircuit. Those RMNAMs are then recombined with the MNAM of the remaining circuit and the reduction process is repeated on the resulting matrix.

To further illustrate the SoE approach, we present the following example.

Example 7. Consider a bipolar cascade stage with bootstrap capacitor C_B illustrated in Figure 8.12 [18]. With the BJTs replaced by their low-frequency hybrid-π models (with r_B, g_m, and r_o only), the full symbolic analysis yields the output admittance formula outlined in Figure 8.13. The formula requires 48 additions and 117 multiplication/division operations. STAINS can generate several different sequences of expressions. One of them is presented in Figure 8.14. It requires only 24 additions and 17 multiplications/divisions.[7]

[4]The subcircuits have to be truly identical, i.e., they must have the same topology and component symbols. A typical example would be a large active filter containing a number of identical, nonideal op amps.

[5]In sensitivity calculations using SoE [5], the final RMNAM may need to be larger than 2×2.

[6]The internal variables are the variables not directly associated with the subcircuit's connections to the rest of the circuit.

[7]Counting of visible arithmetic operations gives only a rough estimate of the SoE complexity, especially when complex numbers are involved. Issues related to SoE computational efficiency are discussed in Reference [55].

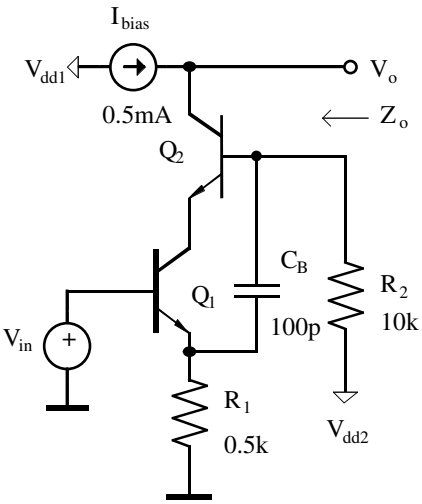


FIGURE 8.12 Bipolar cascode stage.

```
Zo = (G2*Gm1*Gm2+G1*G2*Gm1+G2*Gm2*Gp1+G2*Gm1*Gp2+Gm2*Go1*Gp1+G1*G2*Gp1+G2*Gm2*Go1+...
    G2*Gm1*Go2+G1*Go2*Gp1+G2*Gp1*Gp2+G1*Go1*Gp1+G1*G2*Go2+G1*G2*Go1+Go2*Gp1*Gp2+...
    Go1*Gp1*Gp2+G2*Go2*Gp1+G2*Go1*Gp2+G2*Go2*Gp2+Go1*Go2*Gp1+G2*Go1*Go2+...
    s*(Cb*Gm1*Gm2+Cb*G1*Gm1+Cb*Gm1*Gp2+Cb*G2*Gm1+Cb*G1*Gp1+Cb*Gm2*Go1+Cb*Gp1*Gp2+...
    Cb*G1*Go2+Cb*G2*Gp1+Cb*G1*Go1+Cb*Go1*Gp2+Cb*Go2*Gp2+Cb*Go1*Gp1+Cb*G2*Go1+...
    Cb*G2*Go2+Cb*Go1*Go2))/...
    (Go1*G2*Gm2*Gp1+Go1*G1*G2*Gp1+Go1*G2*Gp1*Gp2+Go1*G1*Go2*Gp1+Go1*G1*G2*Go2+...
    Go1*Go2*Gp1*Gp2+Go1*G2*Go2*Gp2+Go1*G2*Gp2*Gp1+...
    s*(Cb*Go1*G1*Gp1+Cb*Go1*Gp1*Gp2+Cb*Go1*G2*Gp1+Cb*Go1*G1*Go2+Cb*Go1*Go2*Gp2+...
    Cb*Go1*G2*Go2));
```

FIGURE 8.13 Full symbolic expression for Z_o of the cascode in Figure 8.12.

```
d1 = -(G2+Gp2+s*Cb)/(s*Cb);
x1 = (Go1+Gm1)*d1-Gp2-Gm2;
x2 = -s*Cb-(G1+Gp1+Go1+Gm1+s*Cb)*d1;
d2 = Gp2/(s*Cb);
x3 = Go1+Gp2+Go2+Gm2+(Go1+Gm1)*d2;
x4 = -Go1-(G1+Gp1+Go1+Gm1+s*Cb)*d2;
d3 = x2/(x4);
x5 = Gm2+(Go2+Gm2)*d3;
x6 = x1-x3*d3;
Yo = Go2+x5*Go2/(x6);
Zo = 1/Yo;
```

FIGURE 8.14 The SoE generated by STAINS for the cascode in Figure 8.12.

8.5 Approximate Symbolic Analysis

The SoE approach offers a solution for the exact symbolic analysis of large circuits. For some applications, it may be more important to obtain a simpler inexact expression, but the one that would clearly identify the dominant circuit components and their role in determining circuit behavior. Approximate symbolic

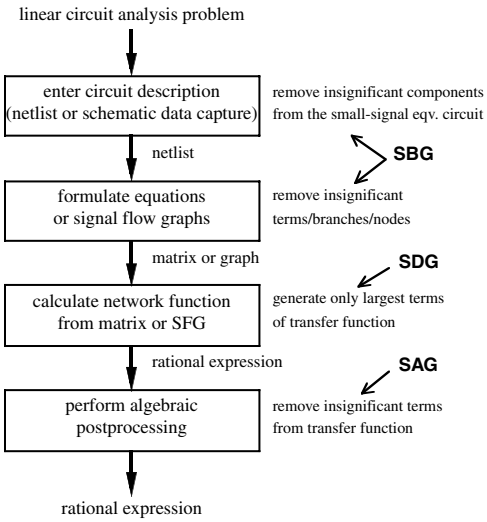


FIGURE 8.15 Classification of symbolic approximation techniques [26].

analysis provides the answer. Of course, manual approximation (simplification) techniques have been known and practiced by engineers for decades. To obtain compact and meaningful expressions by computer, symbolic analysis software must be capable of performing those approximations that are applied in manual circuit analysis in an automatic fashion. In addition to that, computer algorithms should be able to employ simplification strategies not available (or impractical) in manual approximation.

In the last decade, a number of symbolic approximation algorithms have been developed and implemented in symbolic circuit analysis programs. Depending on the stage in the circuit analysis process in which they are applied, these algorithms can be categorized as: *simplification before generation* (SBG), *simplification during generation* (SDG), and *simplification after generation* (SAG). Figure 8.15, adapted from [26], presents an overview of the three types of approximation algorithms.

SBG involves removing circuit components and/or individual entries in the circuit matrix (the *sifting* approach [28]) or eliminating some graph branches (the *sensitivity-based two-graph simplification* [69]) that do not contribute significantly to the final formula.

SDG is based on generation of symbolic terms in a decreasing order of magnitude. The generation process is stopped when the error reaches the specified level. The most successful approach to date is based on the two-graph formulation [68]. It employs an algorithm to generate the common spanning trees in strictly decreasing order of magnitude [30]. In the case of frequency-dependent circuit, this procedure is applied separately to different powers of s. Mathematical formalism of *matroids* is well suited to describe problems of SDG [69].

When applied alone, SAG is a very ineffective technique, because it requires generation and storage of a large number of unnecessary terms. When combined with SBG and SDG methods, however, it can produce the most compact expressions by pruning redundant terms not detected earlier in the simplification process.

All simplification techniques require careful monitoring of the approximation amplitude and phase errors (ε_A and ε_p). The error criteria can be expressed as follows:

$$\left|\frac{\left|H(s,\mathbf{x})\right|-\left|H^*(s,\mathbf{x})\right|}{\left|H(s,\mathbf{x})\right|}\right| \le \varepsilon_A$$

$$\left|\angle H(s,\mathbf{x})-\angle H^*(s,\mathbf{x})\right| \le \varepsilon_p$$

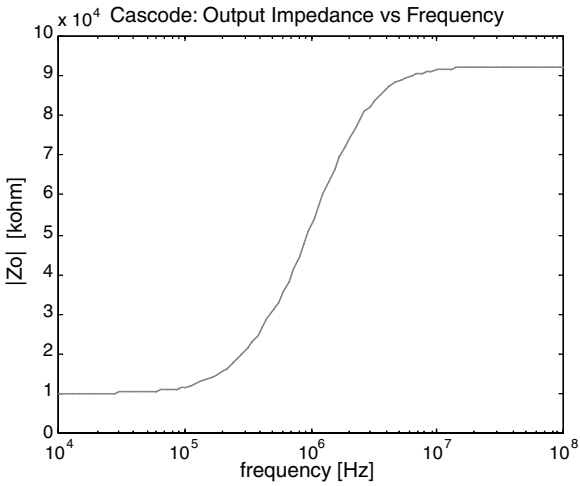


FIGURE 8.16 Plot of $|Z_o|$ of the cascode, obtained numerically from the exact formula.

for $s = j\omega$, $\omega \in (\omega_1,\omega_2)$, and $\mathbf{x} \in (\mathbf{x}_1,\mathbf{x}_2)$, where $H(s,\mathbf{x})$ is the exact transfer function, defined by Eq. (8.3), and $H^*(s,\mathbf{x})$ is the approximating function. The majority of the approximation methods developed to date use the simplified criteria, where the errors are measured only for a given set of circuit parameters $\mathbf{x}^0$ (the nominal design point) [33].

The following example, although quite simple, illustrates very well the advantages of approximate symbolic analysis.

Example 8 [18]. Consider again the bipolar cascode stage, depicted in Figure 8.12 and its fully symbolic expression for the output impedance, depicted in Figure 8.13. Even for such a simple circuit, the full symbolic result is very hard to interpret and therefore not able to provide insight into the circuit behavior. Sequential form of the output impedance formula, presented in Figure 8.14, is more compact than the full expression but also cannot be utilized for interpretation.

A plot of $|Z_o|$ for a nominal set of component values (r_π = 5 kΩ, g_m = 20 mS, r_o = 100 kΩ for both BJTs), obtained numerically from the SoE in Figure 8.14, is plotted in Figure 8.16. By examining the plot, one can appreciate the general behavior of the function, but it is difficult to predict the influence of various circuit components on the output impedance.

Applying symbolic approximation techniques we can obtain less accurate but still more revealing formulas. If a 10% maximum amplitude error is accepted, the simplified function takes the following form:

$$Z_{o(10\%)} = \frac{g_{m1}\left(g_{m2}+G_1\right)\left(G_2+sC_B\right)}{g_{o1}g_{\pi1}\left[G_2\left(g_{m2}+G_1\right)+sC_B\left(G_1+g_{\pi2}\right)\right]}$$

If we allow a 25% magnitude error,[8] the output impedance formula can be simplified further:

$$Z_{o(25\%)} = \frac{g_{m1}g_{m2}\left(G_2+sC_B\right)}{g_{o1}g_{\pi1}\left(G_2g_{m2}+sC_BG_1\right)} \tag{8.25}$$

[8]It is important to note that the approximate expressions were developed taking into account variations of BJT parameters; the fact that both simplified formulas give identical results at the nominal design point is purely coincidental.

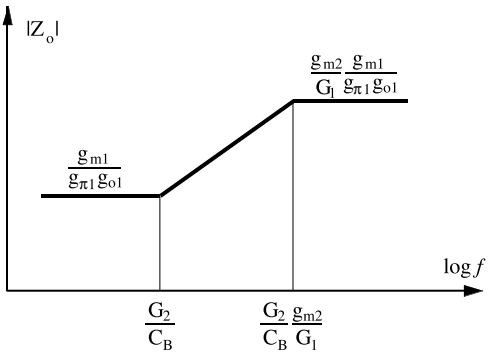


FIGURE 8.17 Asymptotic plot of $|Z_o|$ of the cascode based on Eq. (8.26).

The impedance levels as well as pole and zero estimates can be easily obtained from Eq. (8.25):

$$
\begin{aligned}
Z_o(\text{low } f) &\cong \frac{g_{m1}}{g_{\pi 1} g_{o1}} = \frac{\beta_1}{g_{o1}} \\
Z_o(\text{high } f) &\cong \frac{g_{m1} g_{m2}}{g_{\pi 1} g_{o1} G_1} = \frac{g_{m2}}{G_1} Z_o(\text{low } f) \\
z &\cong -\frac{G_2}{C_B} \\
p &\cong -\frac{g_{m2} G_2}{G_1 C_B}
\end{aligned}
\tag{8.26}
$$

An asymptotic plot of $|Z_o|$, based on Eq. (8.26), is plotted in Figure 8.16.

8.6 Time-Domain Analysis

The previous sections discussed the different frequency domain techniques for symbolic analysis. Symbolic analysis methods in the transient domain did not appear until the beginning of the 1990s [3, 24, 36]. The main limitation to symbolic time-domain analysis is the difficulty in handling the symbolic integration and differentiation needed to handle the energy storage elements (mainly capacitors and inductors). This problem, of course, does not exist in the frequency domain because of the use of Laplace transforms to represent these elements. Although symbolic algebra software packages are available, such as MATHMATICA, MAXIMA, and MAPLE, which can be used to perform integration and differentiations, they have not been applied to transient symbolic analysis due to the execution time complexity of these programs. All but one of the approaches in the time domain are actually semi-symbolic. The semi-symbolic algorithms use a mixture of symbolic and numeric techniques to perform the analysis. The work here is still in its infancy. This section briefly discusses the three contributions published in the literature thus far.

All symbolic time domain techniques deal with linear circuits and can be classified under one of the two categories.

Fully Symbolic

Only one method has been reported in the literature that is fully symbolic [20]. This method utilizes a direct and hierarchical symbolic transient analysis approach similar to the one reported in [22]. The formulation is based on the well-known discrete models for numerical integration of linear differential

equations. Three of these integration methods are implemented symbolically: the Backward Euler, the Trapezoidal, and Gear's 2nd-Order Backward Differentiation [20]. The inherent accuracy problems due to the approximations in these methods show up when the symbolic expressions are evaluated numerically. A detailed discussion of this method can be found in [20].

Semi-Symbolic

Three such algorithms have been reported in the literature thus far. Two of them [24, 36] simply take the symbolic expressions in the frequency domain, evaluate them numerically for a range of frequencies, and then perform a numeric inverse laplace transformation or a fast Fourier transformation (FFT) on the results. The approach reported in [36] uses an MNA, then a state-variable symbolic formulation to get the frequency domain response and can handle time-varying circuits, namely, switch power converters. The approach in [24] uses a hierarchical network approach [22] to generate the symbolic frequency domain response. The third algorithm reported in [3] is a hierarchical approach that uses an MNA and a state-variable symbolic formulation and then uses the eigenvalues of the system to find a closed-form numerical transient solution.

References

[1] G. E. Alderson, P. M. Lin, "Integrating Topological and Numerical Methods for Semi-Symbolic Network Analysis," *Proc. of the 13th Midwest Symposium on Circuit Theory*, 1970.

[2] G. E. Alderson, P. M. Lin, "Computer Generation of Symbolic Network Functions — A New Theory and Implementation," *IEEE Trans. on Circuit Theory*, vol. CT-20, pp. 48–56, Jan. 1973.

[3] B. Alspaugh, M. Hassoun, "A Mixed Symbolic and Numeric Method for Closed-Form Transient Analysis," *Proc. ECCTD*, Davos, 1993.

[4] Z. Arnautovic, P. M. Lin, "Symbolic Analysis of Mixed Continuous and Sampled Data Systems," *Proc. IEEE ISCAS*, pp. 798–801, 1991.

[5] F. Balik, B. Rodanski, "Calculation of First-Order Symbolic Sensitivities in Sequential Form via the Transimpedance Method," *Proc. SMACD*, Kaiserslautern, Germany, Oct. 1998, pp. 169–172.

[6] D. A. Calahan, "Linear Network Analysis and Realization — Digital Computer Programs and Instruction Manual," University of Ill. Bull., vol. 62, Feb. 1965.

[7] L. O. Chua, P. M. Lin, *Computer-Aided Analysis of Electronic Circuits — Algorithms and Computational Techniques*. Englewood Cliffs, NJ: Prentice Hall, 1975.

[8] C. L. Coates, "Flow graph Solutions of Linear Algebraic Equations," *IRE Trans. on Circuit Theory*, vol. CT-6, pp. 170–187, 1959.

[9] F. Constantinescu, M. Nitescu, "Computation of Symbolic Pole/Zero Expressions for Analog Circuit Design," *Proc. SMACD*, Haverlee, Belgium, Oct. 1996.

[10] G. DiDomenico et al., "BRAINS: A Symbolic Solver for Electronic Circuits," *Proc. SMACD*, Paris, Oct. 1991.

[11] G. Dröge, E. H. Horneber, "Symbolic Calculation of Poles and Zeros," *Proc. SMACD*, Haverlee, Belgium, Oct. 1996.

[12] F. V. Fernandez, A. Rodriguez-Vazquez, J. L. Huertas, "An Advanced Symbolic Analyzer for the Automatic Generation of Analog Circuit Design Equations," *Proc. IEEE ISCAS*, Singapore, pp. 810–813, June 1991.

[13] F. V. Fernandez *et al.*, "On Simplification Techniques for Symbolic Analysis of Analog Integrated Circuits," *Proc. IEEE ISCAS*, San Diego, CA, pp. 1149–1152, May 1992.

[14] F. V. Fernandez *et al.*, "Symbolic Analysis of Large Analog Integrated Circuits: The Numerical Reference Generation Problem," *IEEE Trans. on Circuits and Systems — II: Analog and Digital Signal Processing*, vol. 45, no. 10, pp. 1351–1361, Oct. 1998.

[15] J. K. Fidler, J. I. Sewell, "Symbolic Analysis for Computer-Aided Circuit Design — The Interpolative Approach," *IEEE Trans. on Circuit Theory*, vol. CT-20, Nov. 1973.

[16] T. F. Gatts, N. R. Malik, "Topoloigical Analysis Program For Linear Active Networks (TAPLAN)," *Proc. of the 13th Midwest Symposium on Circuit Theory*, 1970.
[17] G. Gielen, H. Walscharts, W. Sansen, "ISSAC: A Symbolic Simulator for Analog Integrated Circuits," *IEEE J. of Solid-State Circuits*, vol. SC-24, pp. 1587–1597, Dec. 1989.
[18] G. Gielen, W. Sansen, *Symbolic Analysis for Automated Design of Analog Integrated Circuits*. Boston, MA: Kluwer Academic, 1991.
[19] S. Greenfield, *Transient Analysis for Symbolic Simulation*, MS Thesis, Iowa State University, Dec. 1993.
[20] S. Greenfield, M. Hassoun, "Direct Hierarchical Symbolic Transient Analysis of Linear Circuits," *Proc. ISCAS*, 1994.
[21] G. D. Hachtel et al., "The Sparse Tableau Approach to Network and Design," *IEEE Trans. on Circuit Theory*, vol. CT-18, pp. 101–113, Jan 1971.
[22] M. M. Hassoun, P. M. Lin, "A New Network Approach to Symbolic Simulation of Large-Scale Networks," *Proc. IEEE ISCAS*, pp. 806–809, May 1989.
[23] M. M. Hassoun, P. M. Lin, "An Efficient Partitioning Algorithm for Large-Scale Circuits," Proc. *IEEE ISCAS*, New Orleans, pp. 2405–2408, May 1990.
[24] M. M. Hassoun, J. E. Ackerman, "Symbolic Simulation of Large Scale Circuits in Both Frequency and Time Domains," *Proc. IEEE MWSCAS*, Calgary, pp. 707–710, Aug. 1990.
[25] M. Hassoun, K. McCarville, "Symbolic Analysis of Large-Scale Networks Using a Hierarchical Signal Flow Graph Approach," *J. of Analog VLSI and Signal Processing*, Jan. 1993.
[26] E. Henning, Symbolic Approximation and Modeling Techniques for Analysis and Design of Analog Circuits. Doctoral Dissertation, University of Kaiserslautern. Aachen: Shaker Verlag, 2000.
[27] C. Ho, A. E. Ruehli, P. A. Brennan, "The Modified Nodal Approach to Network Analysis," *IEEE Trans. on Circuits and Systems*, vol. CAS-25, pp. 504–509, June 1975.
[28] J. J. Hsu, C. Sechen, "Low-Frequency Symbolic Analysis of Large Analog Integrated Circuits," *Proc. CICC*, 1993, pp. 14.7.1–14.7.4.
[29] L. Huelsman, "Personal Computer Symbolic Analysis Programs for Undergraduate Engineering Courses," *Proc. ISCAS*, pp. 798–801, 1989.
[30] N. Katoh, T. Ibaraki, H. Mine, "An Algorithm for Finding k Minimum Spanning Trees," *SIAM J. Comput.*, vol. 10, no. 2, pp. 247–255, May 1981.
[31] A. Konczykowska, M. Bon, "Automated Design Software for Switched Capacitor ICs with Symbolic Simulator SCYMBAL," *Proc. DAC*, pp. 363–368, 1988.
[32] A. Konczykowska et al., "Symbolic Analysis as a Tool for Circuit Optimization," *Proc. IEEE ISCAS*, San Diego, CA, pp. 1161–1164, May 1992.
[33] A. Konczykowska, "Symbolic circuit analysis," in *Wiley Encyclopedia of Electrical and Electronics Engineering*, J. G. Webster, Ed. New York: John Wiley & Sons, 1999.
[34] J. Lee, R. Rohrer, "AWEsymbolic: Compiled Analysis of Linear(ized) Circuits Using Asymptotic Waveform Evaluation," *Proc. DAC*, pp. 213-218, 1992.
[35] B. Li, D. Gu, "SSCNAP: A Program for Symbolic Analysis of Switched Capacitor Circuits," *IEEE Trans. on CAD*, vol. 11, pp. 334–340, 1992.
[36] A. Liberatore et al., "Simulation of Switching Power Converters Using Symbolic Techniques," *Alt Frequenza*, vol. 5, no. 6, Nov. 1993.
[37] A. Liberatore et al., "A New Symbolic Program Package for the Interactive Design of Analog Circuits," *Proc. IEEE ISCAS*, Seattle, WA, pp. 2209–2212, May 1995.
[38] P. M. Lin, G. E. Alderson, "SNAP — A Computer Program for Generating Symbolic Network Functions," School of EE, Purdue University, West Lafayette, IN, Rep. TR-EE 70-16, Aug. 1970.
[39] P. M. Lin, "A Survey of Applications of Symbolic Network Functions," *IEEE Trans. on Circuit Theory*, vol. CT-20, pp. 732–737, Nov. 1973.
[40] P. M. Lin, *Symbolic Network Analysis*. Amsterdam: Elsevier Science, 1991.
[41] P. M. Lin, "Sensitivity Analysis of Large Linear Networks Using Symbolic Programs," *Proc. IEEE ISCAS*, San Diego, CA, pp. 1145–1148, May 1992.

[42] V. K. Manaktala, G. L. Kelly, "On the Symbolic Analysis of Electrical Networks," *Proc. of the 15th Midwest Symposium on Circuit Theory*, 1972.
[43] S. Manetti, "New Approaches to Automatic Symbolic Analysis of Electric Circuits," *Proc. IEE*, pp. 22–28, Feb. 1991.
[44] M. Martins et al., "A Computer-Assisted Tool for the Analysis of Multirate SC Networks by Symbolic Signal Flow Graphs," *Alt Frequenza*, vol. 5, no. 6, Nov. 1993.
[45] S. J. Mason, "Feedback Theory — Further Properties of Signal Flow Graphs," *Proc. IRE*, vol. 44, pp. 920–926, July 1956.
[46] J. O. McClanahan, S. P. Chan, "Computer Analysis of General Linear Networks Using Digraphs," *Int. J. of Electronics*, no. 22, pp. 153–191, 1972.
[47] L.P. McNamee, H. Potash, *A User's and Programmer's Manual for NASAP*, University of California at Los Angeles, Rep. 63-38, Aug. 1968.
[48] R. R. Mielke, "A New Signal Flowgraph Formulation of Symbolic Network Functions," *IEEE Trans. on Circuits and Systems*, vol. CAS-25, pp. 334–340, June 1978.
[49] H. Okrent, L. P. McNamee, *NASAP-70 User's and Programmer's Manual*, UCLA, Technical Report ENG-7044, 1970.
[50] M. Pierzchala, B. Rodanski, "A New Method of Semi-Symbolic Network Analysis," *Proc. IEEE ISCAS*, Chicago, IL, pp. 2240–2243, May 1993.
[51] M. Pierzchala, B. Rodanski, "Efficient Generation of Symbolic Network Functions for Large-Scale Circuits," *Proc. MWSCAS*, Ames, IO, pp. 425–428, August 1996.
[52] M. Pierzchala, B. Rodanski, "Direct Calculation of Numerical Coefficients in Semi-Symbolic Circuit Analysis," *Proc. SMACD*, Kaiserslautern, Germany, Oct. 1998, pp. 173–176.
[53] M. Pierzchala, B. Rodanski, "Generation of Sequential Symbolic Network Functions for Large-Scale Networks by Circuit Reduction to a Two-Port," *IEEE Trans. on Circuits and Systems — I: Fundamental Theory and Applications*, vol. 48, no. 7, July 2001.
[54] C. Pottle, *CORNAP User Manual*, School of Electrical Engineering, Cornell University, Ithaca, NY, 1968.
[55] B. Rodanski, "Computational Efficiency of Symbolic Sequential Formulae," *Proc. SMACD*, Lisbon, Portugal, pp. 45-50, Oct. 2000.
[56] P. Sannuti, N. N. Puri, "Symbolic Network Analysis — An Algebraic Formulation," *IEEE Trans. on Circuits and Systems*, vol. CAS-27, pp. 679–687, Aug. 1980.
[57] S. Seda, M. Degrauwe, W. Fichtner, "Lazy-Expansion Symbolic Expression Approximation in SYNAP," *1992 Int. Conf. Computer-Aided Design,* Santa Clara, CA, pp. 310–317, 1992.
[58] K. Singhal, J. Vlach, "Generation of Immittance Functions in Symbolic Form for Lumped Distributed Active Networks," *IEEE Trans. on Circuits and Systems*, vol. CAS-21, pp. 57–67, Jan. 1974.
[59] K. Singhal, J. Vlach, "Symbolic Analysis of Analog and Digital Circuits," *IEEE Trans. on Circuits and Systems*, vol. CAS-24, pp. 598–609, Nov. 1977.
[60] R. Sommer, "EASY — An Experimental Analog Design System Framework," *Proc. SMACD*, Paris, Oct. 1991.
[61] J. A. Starzyk, A. Konczykowska, "Flowgraph Analysis of Large Electronic Networks," *IEEE Trans. on Circuits and Systems*, vol. CAS-33, pp. 302–315, March 1986.
[62] J. A. Starzyk, J. Zou "Direct Symbolic Analysis of Large Analog Networks," *Proc. MWSCAS*, Ames, IO, pp. 421–424, Aug. 1996.
[63] M. D. Topa, "On Symbolic Analysis of Weakly-Nonlinear Circuits," *Proc. SMACD*, Kaiserslautern, Germany, Oct. 1998, pp. 207–210.
[64] J. Vlach, K. Singhal, *Computer Methods for Circuit Analysis and Design*, 2nd ed. New York: Van Nostrand Reinhold, 1994.
[65] C. Wen, H. Floberg, Q. Shui-sheng, "A Unified Symbolic Method for Steady-State Analysis of Nonlinear Circuits and Systems," *Proc. SMACD*, Kaiserslautern, Germany, Oct. 1998, pp. 218–222.

[66] G. Wierzba et al., "SSPICE — A Symbolic SPICE Program for Linear Active Circuits," *Proc. MWSCAS*, 1989.

[67] P. Wambacq, G. Gielen, W. Sansen, "A Cancellation-Free Algorithm for the Symbolic Simulation of Large Analog Circuits," *Proc. ISCAS*, San Diego, CA, pp. 1157–1160, May 1992.

[68] P. Wambacq, G. E. Gielen, W. Sansen, "Symbolic Network Analysis Methods for Practical Analog Integrated Circuits: A Survey," *IEEE Trans. on Circuits and Systems — II: Analog and Digital Signal Processing*, vol. 45, no. 10, pp. 1331–1341, Oct. 1998.

[69] Q. Yu, C. Sechen, "A Unified Approach to the Approximate Symbolic Analysis of Large Analog Integrated Circuits," *IEEE Trans. on Circuits and Systems — I: Fundamental Theory and Applications*, vol. 43, no. 8, pp. 656–669, Aug. 1996.

9

Analysis in the Time Domain

Robert W. Newcomb
University of Maryland

9.1 Signal Types

Introduction

Because information into and out of a circuit is carried via time domain signals we look first at some of the basic signals used in continuous time circuits. All signals are taken to depend on continuous time t over the full range $-\infty < t < \infty$. It is important to realize that not all signals of interest are functions in the strict mathematical sense; we must go beyond them to generalized functions (e.g., the impulse), which play a very important part in the signal processing theory of circuits.

Step, Impulse, and Ramp

The unit **step** function, denoted $1(\cdot)$, characterizes sudden jumps, such as when a signal is turned on or a switch is thrown; it can be used to form pulses, to select portions of other functions, and to define the ramp and impulse as its integral and derivative. The unit step function is discontinuous and jumps between two values, 0 and 1, with the time of jump between the two taken as $t = 0$. Precisely,

$$1(t) = \begin{cases} 1 & \text{if } t > 0 \\ 0 & \text{if } t < 0 \end{cases} \tag{9.1}$$

which is illustrated in Figure 9.1 along with some of the functions to follow.

Here, the value at the jump point, $t = 0$, purposely has been left free because normally it is immaterial and specifying it can lead to paradoxical results. Physical step functions used in the laboratory are actually continuous functions that have a continuous rise between 0 and 1, which occurs over a very short time. Nevertheless, instances occur in which one may wish to set $1(0)$ equal to 0 or to 1 or to 1/2 (the latter, for example, when calculating the values of a Fourier series at a discontinuity). By shifting the time

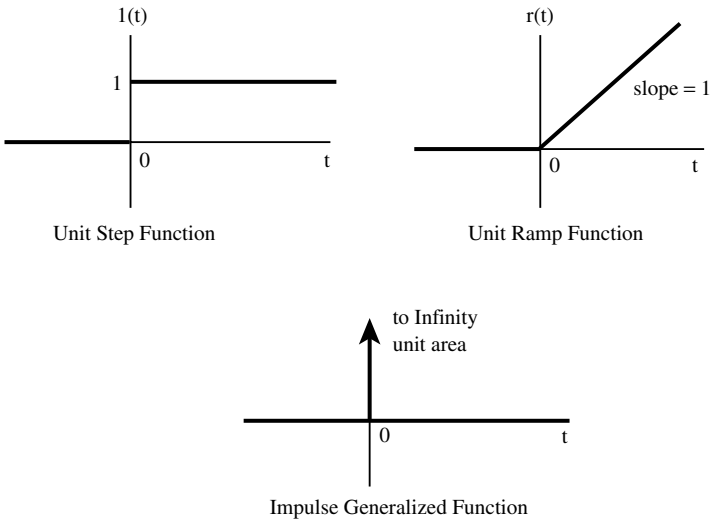


FIGURE 9.1 Step, ramp, and impulse functions.

argument the jump can be made to occur at any time, and by multiplying by a factor the height can be changed. For example, $1(t - t_0)$ has a jump at time t_0 and $a[1(t) - 1(t - t_0)]$ is a pulse of width t_0 and height a going up to a at $t = 0$ and down to 0 at time t_0. If $a = a(t)$ is a function of time, then that portion of $a(t)$ between 0 and t_0 is selected. The unit **ramp**, $r(\cdot)$ is the continuous function which ramps up linearly (with unit slope) from zero starting at $t = 0$; the ramp results from the unit step by integration

$$r(t) = \int_{-\infty}^{t} 1(\tau)d\tau = t1(t) = \begin{cases} t & \text{if } t > 0 \\ 0 & \text{if } t < 0 \end{cases} \tag{9.2}$$

As a consequence the unit step is the derivative of the unit ramp, while differentiating the unit step yields the unit **impulse** generalized function, $\delta(\cdot)$ that is

$$\delta(t) = \frac{d^1(t)}{dt} = \frac{d^2 r(t)}{dt^2} \tag{9.3}$$

In other words, the unit impulse is such that its integral is the unit step; that is, its area at the origin, $t = 0$, is 1. The impulse acts to sample continuous functions which multiply it, i.e.,

$$a(t)\delta(t - t_0) = a(t_0)\delta(t - t_0) \tag{9.4}$$

This sampling property yields an important integral representation of a signal $x(\cdot)$

$$\begin{aligned} x(t) &= \int_{-\infty}^{\infty} x(\tau)\delta(t - \tau)d\tau \\ &= \int_{-\infty}^{\infty} x(t)\delta(t - \tau)d\tau = x(t) \int_{-\infty}^{\infty} \delta(t - \tau)d\tau \end{aligned} \tag{9.5}$$

where the validity of the first line is seen from the second line, and the fact that the integral of the impulse through its jump point is unity. Equation (9.5) is actually valid even when $x(\cdot)$ is discontinuous and, consequently, is a fundamental equation for linear circuit theory. Differentiating $\delta(t)$ yields an even more discontinuous object, the doublet $\delta'(\cdot)$. Strictly speaking, the impulse, all its derivatives, and signals of

that class are not functions in the classical sense, but rather they are operators [1] or functionals [2], called generalized functions or, often, distributions. Their evaluations take place via test functions, just as voltages are evaluated on test meters.

The importance of the impulse lies in the fact that if a linear time-invariant system is excited by the unit impulse, then the response, naturally called the impulse response, is the inverse Laplace transform of the network function. In fact, if $h(t)$ is the impulse response of a linear time-invariant (continuous and continuous time) circuit, the forced response $y(t)$ to any input $u(t)$ can be obtained without leaving the time domain by use of the convolution integral, with the operation of convolution denoted by $*$,

$$y(t) = h * u = \int_{-\infty}^{\infty} h(t-\tau)u(\tau)d\tau \tag{9.6}$$

Equation (9.6) is mathematically rigorous, but justified on physical grounds through (9.5) as follows. If we let $h(t)$ be the output when $\delta(t)$ is the input, then, by time invariance, $h(t-\tau)$ is the output when the input is shifted to $\delta(t-\tau)$. Scaling the latter by $u(\tau)$ and summing via the integral, as designated in (9.5), we obtain a representation of the input $u(t)$. This must result in the output representation being in the form of (9.6) by linearity of the system through similar scaling and summing of $h(t-\tau)$, as was performed on the input.

Sinusoids

Sinusoidal signals are important because they are self-reproducing functions (i.e., eigenfunctions) of linear time-invariant circuits. This is true basically because the derivatives of sinusoids are sinusoidal. As such, sinusoids are also the natural outputs of oscillators and are delivered in power sources, including laboratory signal generators and electricity for the home derived from the power company.

Eternal

Eternal signals are defined as being of the same nature for all time, $-\infty < t < \infty$, in which case an eternal cosine repeats itself eternally in both directions of time, with an origin of time, $t = 0$, being arbitrarily fixed. Because eternal sinusoids have been turned on forever, they are useful in describing the steady operation of circuits. In particular, the signal $A\cos(\omega t + \theta)$ over $-\infty < t < \infty$ defines an eternal cosine of amplitude A, radian frequency $\omega = 2\pi f$ (with f being real frequency, in Hertz, which are cycles per second), at phase angle θ (in radians and with respect to the origin of time), with A, ω, and θ real numbers. When $\theta = \pi/2$ this cosine also represents a sine, so that all eternal sinusoidal signals are contained in the expression $A\cos(\omega t + \theta)$.

At times, it is important to work with sinusoids that have an exponential envelope, with the possibility that the envelope increases or decreases with time, that is, with positively or negatively damped sinusoids. These are described by $Ae^{st}\cos(\omega t + \theta)$, where the real number is the damping factor, giving signals that damp out in time when the damping factor is positive and signals that increase with time when the damping factor is negative. Of most importance when working with this class of signals is the identity

$$e^{\sigma t + j\omega t} = e^{st} = e^{\sigma t}\left[\cos(\omega t) + j\sin(\omega t)\right] \tag{9.7}$$

where $s = \sigma + j\omega$ with $j = \sqrt{-1}$. Here, s is called the **complex frequency**, with its imaginary part being the real (radian) frequency, ω. When no damping is present, $s = j\omega$, in which case the exponential form of (9.7) represents pure sinusoids. In fact, we see in this expression that the cosine is the real part of an exponential and the sine is its imaginary part. Because exponentials are usually easier than sinusoids to treat analytically, the consequence for real linear networks is that we can do most of the calculations with exponentials and convert back to sinusoids at the end. In other words, if a real linear system has a cosine or a damped cosine as a true input, it can be analyzed by using instead the exponential of which it is the real part as its (fictitious) input, finding the resulting (fictitious) exponential output, and then taking

the real part at the end of the calculations to obtain the true output for the true input. Because exponentials are probably the easiest signals to work with in theory, the use of exponentials rather than sinusoids usually greatly simplifies the theory and calculations for circuits operating under steady-state conditions.

Causal

Because practical circuits have not existed since $t = -\infty$ they usually begin to be considered at a suitable starting time, taken to be $t = 0$, in which case the associated signals can be considered to be zero for $t < 0$. Mathematically, these functions are said to have support bounded on the left. The support of a signal is (the closure of) that set of times for which the signal is non-zero, therefore, the support of these signals is bounded on the left by zero. When signals are discontinuous functions they have the important property that they can be represented by multiplying with unit step functions signals which are differentiable and have nonbounded support. For example, $g(t) = e^{st} \cdot 1(t)$ has a jump at $t = 0$ with support at the half line 0 to ∞ but has e^{st} infinitely differential of "eternal" support.

A **causal** circuit is one for which the response is only nonzero after the input becomes nonzero. Thus, if the inputs are zero for $t < 0$, the outputs of causal circuits are also zero for $t < 0$. In such cases the impulse response, $h(t)$, or the response to an input impulse of "infinite jump" at $t = 0$, satisfies $h(t) = 0$ for $t < 0$ and the convolution form of the output, (9.4), takes the form

$$y(t) = \left[\int_0^t h(t-\tau)u(\tau)d\tau\right]1(t) \tag{9.8}$$

Periodic and Aperiodic Waveforms

The pure sinusoids, although not the sinusoids with nonzero damping, are special cases of periodic signals. In other words, ones which repeat themselves in time every T seconds, where T is the period. Precisely, a time-domain signal $g(\cdot)$ is **periodic** of period T if $g(t) = g(t + T)$, where normally T is taken to be the smallest nonzero T for which this is true. In the case of the sinusoids, $A\cos(\omega t + \theta)$ with $\omega = 2\pi f$, the period is given by $T = 1/f$ because $\{2\pi[f(t + T)] + \theta\} = \{2\pi ft + 2\pi(fT) + \theta\} = \{2\pi ft + (2\pi + \theta)\}$, and sinusoids are unchanged by a change of 2π in the phase angle. Periodic signals need to be specified over only one period of time, e.g., $0 \le t < T$, and then can be extended periodically for all time by using $t = t \bmod(T)$ where $\bmod(\cdot)$ is the modulus function; in other words, periodic signals can be looked upon as being defined on a circle, if we imagine the circle as being a clock face.

Periodic signals represent rhythms of a system and, as such, contain recurring information. As many phycial systems, especially biomedical systems, either possess directly or to a very good approximation such rhythms, the periodic signals are of considerable importance. Even though countless periodic signals are available besides the sinusoids, it is important to note that almost all can be represented by a Fourier series. Exponentials are eigenfunctions for linear circuits, thus, the Fourier series is most conveiently expressed for circuit considerations in terms of the exponential form. If $g(t) = g(t + T)$, then

$$g(t) \cong \sum_{n=-\infty}^{\infty} c_n e^{j(2\pi nt/T)} \tag{9.9}$$

where the coefficients are complex and are given by

$$c_n = \frac{1}{T}\int_0^T g(t)e^{-j(2\pi nt/T)}dt = a_n + jb_n \tag{9.10}$$

Strictly speaking, the integral is over the half-open interval $[0,T)$ as seen by considering $g(\cdot)$ defined on the circle. In (9.9), the symbol $\cong$ is used to designate the expression on the right as a representation that

may not exactly agree numerically with the left side at every point when $g(\cdot)$ is a function; for example, at discontinuities the average is obtained on the right side. If $g(\cdot)$ is real, that is, $g(t) = g(t)^*$, where the superscript * denotes complex conjugate, then the complex coefficients c_n satisfy $c_n = c_{-n}^*$. In this case the real coefficients a_n and b_n in (9.10) are even and odd in the indices; n and the a_n combine to give a series in terms of cosines, and the b_n gives a series in terms of sines.

As an example the square wave, sqw(t), can be defined by

$$\mathrm{sqw}(t) = 1(t) - 1\left(t - \lfloor T/2 \rfloor\right) \quad 0 \le t < T \tag{9.11}$$

and then extended periodically to $-\infty < t < \infty$ by taking $t = t \bmod(T)$. The exponential Fourier series coefficients are readily found from (9.10) to be

$$c_n = \begin{cases} 1/2 & \text{if } n = 0 \\ \dfrac{1}{j\pi n} \begin{cases} 0 & \text{if } n = 2k \neq 0 \ (\text{even} \neq 0) \\ 1 & \text{if } n = 2k+1 \ (\text{odd}) \end{cases} \end{cases} \tag{9.12}$$

for which the Fourier series is

$$\mathrm{sqw}(t) \cong \frac{1}{2} + \sum_{k=-\infty}^{\infty} \frac{1}{j\pi[2k+1]} e^{j2\pi[2k+1]t/T} \tag{9.13}$$

The derivative of sqw(t) is a periodic set of impulses

$$\frac{d[\mathrm{sqw}(t)]}{dt} = \delta(t) - \delta\left(t - \lfloor T/2 \rfloor\right) \quad 0 \le t < T \tag{9.13}$$

for which the exponential Fourier series is easily found by differentiating (9.13), or by direct calculation from (9.10), to be

$$\sum_{i=-\infty}^{\infty} \left(\delta(t - iT) - \delta\left(t - iT - \lfloor T/2 \rfloor\right)\right) \cong \sum_{k=-\infty}^{\infty} \frac{2}{T} e^{j\left(2\pi[2k+1]t/T\right)} \tag{9.15}$$

Combining the exponentials allows for a sine representation of the periodic generalized function signal. Further differentiation can take place, and by integrating (9.15) we get the Fourier series for the square wave if the appropriate constant of integration is added to give the DC value of the signal. Likewise, a further integration will yield the Fourier series for the sawtooth periodic signal, and so on.

The importance of these Fourier series representations is that a circuit having periodic signals can always be considered to be processing these signals as exponential signals, which are usually self-reproducing signals for the system, making the design or analysis easy. The Fourier series also allows visualization of which radian frequencies, $2\pi n/T$, may be important to filter out or emphasize. In many common cases, especially for periodically pulsed circuits, the series may be expressed in terms of impulses. Thus, the impulse response of the circuit can be used in conjunction with the Fourier series.

References

[1] J. Mikusinski, *Operational Calculus*, 2nd ed., New York; Pergamon Press, 1983.

[2] A. Zemanian, *Distribution Theory and Transform Analysis*, New York: McGraw-Hill, 1965.

9.2 First-Order Circuits

Introduction

First-order circuits are fundamental to the design of circuits because higher order circuits can be considered to be constructed of them. Here, we limit ourselves to single-input-output linear time-invariant circuits for which we take the definition of a first-order circuit to be one described by the differential equation

$$d_1 \cdot \frac{dy}{dt} + d_0 \cdot y = n_1 \cdot \frac{du}{dt} + n_0 \cdot u \tag{9.16}$$

where d_0 and d_1 are "denominator" constants and n_0 and n_1 are "numerator" constants, $y = y(\cdot)$ is the output and $u = u(\cdot)$ is the input, and both u and y are generalized functions of time t. So that the circuit truly will be first order, we require that $d_1 \cdot n_0 - d_0 \cdot n_1 \neq 0$, which guarantees that at least one of the derivatives is actually present, but if both derivatives occur, the expressions in y and in u are not proportional, which would lead to cancellation, forcing y and u to be constant multiples of each other. Because a factorization of real higher-order systems may lead to complex first-order systems, we will allow the numerator and denominator constants to be complex numbers; thus, y and u may be complex-valued functions.

If the derivative is treated as an operator, $p = d[\cdot]/dt$, then (9.16) can be conveniently written as

$$y = \frac{n_1 p + n_0}{d_1 p + d_0} u = \begin{cases} \left[\frac{n_1}{d_0} p + \frac{n_0}{d_0}\right] u & \text{if } d_1 = 0 \\ \left[\frac{n_1}{d_1} + \frac{d_1 n_0 - d_0 n_1}{p + (d_0/d_1)}\right] u & \text{if } d_1 \neq 0 \end{cases} \tag{9.17}$$

where the two cases in terms of d_1 are of interest because they provide different forms of responses, each of which frequently occurs in first-order circuits. As indicated by (9.17), the transfer function

$$H(p) = \frac{n_1 p + n_0}{d_1 p + d_0} \tag{9.18}$$

is an operator (as a function of the derivative operator p), which characterizes the circuit. Table 9.1 lists some of the more important types of different first-order circuits along with their transfer functions and causal impulse responses.

The following treatment somewhat follows that given in [1], although with a slightly different orientation in order to handle all linear time-invariant continuous time continuous circuits.

Zero-Input and Zero-State Response

The response of a linear circuit is, via the linearity, the sum of two responses, one due to the input when the circuit is initially in the zero state, called the **zero-state response**, and the other due to the initial state when no input is present, the **zero-input response**. By the linearity the total response is the sum of the two separate responses, and thus we may proceed to find each separately. In order to investigate these two types of responses, we introduce the state vector $x(\cdot)$ and the state-space representation (as previously $p = d[\cdot]/dt$)

$$\begin{aligned} px &= Ax + Bu \\ y &= Cx + Du + Epu \end{aligned} \tag{9.19}$$

TABLE 9.1 Typical Transfer Functions of First-Order Circuits

Transfer Function	Description	Impulse Response
$\frac{n_1}{d_0}p$	Differentiator	$\frac{n_1}{d_0}\delta'(t)$
$\frac{n_0}{d_1 p}$	Integrator	$\frac{n_0}{d_1}1(t)$
$\frac{n_1 p + n_0}{d_1}$	Leaky differentiator	$\frac{n_0}{d_1}\delta(t) + \frac{n_1}{d_1}\delta'(t)$
$\frac{n_0}{d_1 p + d_0}$	Low-pass filter; lossy integrator	$\frac{n_0}{d_1}e^{-\frac{d_0}{d_1}t}\cdot 1(t)$
$\frac{n_1 p}{d_1 p + d_0}$	High-pass filter	$\frac{n_1}{d_1}\delta(t) + \frac{n_1 d_0}{d_1^2}e^{-\frac{d_0}{d_1}t}\cdot 1(t)$
$\frac{n_1}{d_1}\frac{p-(d_0/d_1)}{p+(d_0/d_1)}$	All-pass filter	$\frac{n_1}{d_1}\left[\delta(t) - 2\frac{d_0}{d_1}e^{-\frac{d_0}{d_1}t}\cdot 1(t)\right]$

where A, B, C, D, E are constant matrices. For our first-order circuit two cases are exhibited, depending upon d_1 being zero or not. In the case of $d_1 = 0$,

$$y = (n_1/d_0)u + (n_1/d_0)pu \qquad d_1 = 0 \tag{9.20a}$$

Here, $C = 0$ and A and B can be chosen anything, including empty. When $d_1 \neq 0$, our first-order circuit has the following set of (minimal size) state-variable equations

$$\begin{aligned} px &= \left[-\frac{d_0}{d_1}\right]\cdot x + \left[d_1 n_0 - d_0 n_1\right]\cdot u \\ y &= [1]\cdot x + \left[\frac{n_1}{d_1}\right]\cdot u \end{aligned} \qquad d_1 \neq 0 \tag{9.20b}$$

By choosing $u = 0$ in (9.2), we obtain the equations that yield the zero input response. Specifically, the zero-input response is

$$y(t) = \begin{cases} 0 & \text{if } d_1 = 0 \\ e^{-\frac{d_0}{d_1}t}\cdot y(0) & \text{if } d_1 \neq 0 \end{cases} \tag{9.21}$$

which is also true by direct substitution into (9.16). Here, we have set, in the $d_1 \neq 0$ case, the initial value of the state, $x(0)$, equal to the initial value of the output, $y(0)$, which is valid by our choice of state-space equations. Note that (9.21) is valid for all time and y at $t = 0$ assumes the assigned initial value $y(0)$, which must be zero when the input is zero and no derivative occurs on the output.

The zero-state response is explained as the solution of (9.21) when $x(0) = 0$. In the case that $d_1 = 0$, the zero-state response is

$$y = \frac{n_0}{d_0}u + \frac{n_1}{d_0}pu = \left\{\frac{n_0}{d_0}\delta(t) + \frac{n_1}{d_0}\delta'(t)\right\} * u \qquad d_1 = 0 \tag{9.22a}$$

where $*$ denotes convolution, $\delta(\cdot)$ is the unit impulse, and $1(\cdot)$ is the unit step function. While in the case that $d_1 \neq 0$

$$y = \left\{ \frac{n_1}{d_1}\delta(t) + \left[\frac{d_1 n_0 - d_0 n_1}{d_1} \right] e^{-\frac{d_0}{d_1}t} 1(t) \right\} * u \qquad d_1 \neq 0 \tag{9.22b}$$

which is found by eliminating x from (9.20b) and can be checked by direct substitution into (9.16). The terms in the braces are the causal impulse responses, $h(t)$, which are checked by letting $u = \delta$ with otherwise zero initial conditions, that is, with the circuit initially in the zero state. Actually, infinitely many noncausal impulse responses could be used in (9.22b). One such response is found by replacing $1(t)$ by $-1(-t)$]. However, physically the causal responses are of most interest.

If $d_1 \neq 0$, the form of the responses is determined by the constant d_0/d_1, the reciprocal of which (when $d_0 \neq 0$) is called the **time constant**, t_c, of the circuit because the circuit impulse response decays to $1/e$ at time $t_c = d_1/d_0$. If the time constant is positive, the zero-input and the impulse responses asymptotically decay to zero as time approaches positive infinity, and the circuit is said to be **asymptotically stable.** On the other hand, if the time constant is negative, then these two responses grow without bounds as time approaches plus infinity, and the circuit is called unstable. It should be noted that as time goes in the reverse direction to minus infinity, the unstable zero-input response decays to zero. If $d_0/d_1 = 0$ the zero-input and impulse responses are still stable, but neither decay nor grow as time increases beyond zero.

By linearity of the circuit and its state-space equations, the total response is the sum of the zero-state response and the zero-input response; thus, even when $d_0 = 0$ or $d_1 = 0$

$$y(t) = e^{-\frac{d_0}{d_1}} y_0 + h(t) * u(t) \tag{9.23}$$

Assuming that u and h are zero for $t < 0$ their convolution is also zero for $t < 0$, although not necessarily at $t = 0$, where it may even take on impulsive behavior. In such a case, we see that y_0 is the value of the output instantaneously before $t = 0$. If we are interested only in the circuit for $t > 0$, surprisingly, an input will yield the zero input response. That is, an equivalent input u_0 exists, which will yield the zero input response for $t > 0$, this being $u_0(t) = d_1 y_0 \exp(-t d_0/d_1)1(t)$. Thus, $y = h * (u + u_0)$ gives the same result as (9.23).

When $d_1 = 0$, the circuit acts as a differentiator and within the state-space framework it is treated as a special case. However, in practice it is not a special case because the current, i, versus voltage, v, for a capacitor of capacitance C, in parallel with a resistor of conductance G is described by $i = Cpv + Gv$. Consequently, it is worth noting that all cases can be handled identically in the semistate description

$$\begin{bmatrix} d_1 & d_1 - 1 \\ 0 & 0 \end{bmatrix} px = \begin{bmatrix} -d_0 & -d_0 \\ 0 & 1 \end{bmatrix} x + \begin{bmatrix} n_0 \\ n_1 \end{bmatrix} u$$
$$y = \begin{bmatrix} 1 & 1 \end{bmatrix} x \tag{9.24}$$

where $x(\cdot)$ is the semistate instead of the state, although the first components of the two vectors agree in many cases. In other words, the semistate description is more general than the state description, and handles all circuits in a more convenient fashion [2].

Transient and Steady-State Responses

This section considers stable circuits, although the techniques are developed so that they apply to other situations. In the asymptotically stable case, the zero input response decays eventually to zero; that is, transient responses due to initial conditions eventually will not be felt and concentration can be placed

upon the zero-state response. Considering first eternal exponential inputs, $u(t) = U\exp(\mathrm{st})$ for $-\infty < t < \infty$ at the complex frequency $s = \sigma + j\omega$, where s is chosen as different from the natural frequency $s_n = -d_0/d_1 = -1/t_c$ and U is a constant, we note that the response is $y(t) = Y(s)\exp(\mathrm{st})$, as is observed by direct substitution into (9.16); this substitution yields directly

$$Y(s) = \frac{n_1 s + n_0}{d_1 s + d_0} \cdot U \tag{9.25}$$

where $y(t) = Y(s)\exp(\mathrm{st})$ for $u(t) = U\exp(\mathrm{st})$ over $-\infty < t < \infty$. That is, an exponential excitation yields an exponential response at the same (complex) frequency $s = \sigma + j\omega$ as that for the input. When $\sigma = 0$, the excitation and response are both sinusoidal and the resulting response is called the **sinusoidal steady state** (SSS). Equation (9.25) shows that the SSS response is found by substituting the complex frequency $s = j\omega$ into the **transfer function**, now evaluated on complex numbers instead of differential operators as in (9.18),

$$H(s) = \frac{n_1 s + n_0}{d_1 s + d_0} \tag{9.26}$$

This transfer function represents the impulse response, $h(t)$, of which it is actually the Laplace transform, and as we found earlier, the causal impulse response is

$$h(t) = \begin{cases} \dfrac{n_0}{d_0}\delta(t) + \dfrac{n_1}{d_0}\delta'(t), & \text{if } d_1 = 0 \\ \dfrac{n_1}{d_1}\delta(t) + \left[\dfrac{d_1 n_0 - d_0 n_1}{d_1}\right] e^{-\frac{d_0}{d_1}t} 1(t), & \text{if } d_1 \neq 0 \end{cases} \tag{9.27}$$

However, practical signals are started at some finite time, normalized here to $t = 0$, instead of at $t = -\infty$, as used for the preceding exponentials. Thus, consider an input of the same type but applied only for $t > 0$; i.e., let $u(t) = U\exp(\mathrm{st})1(t)$. The output is found by using the convolution $y = h * u$; after a slight amount of calculation is evaluated to

$$\begin{aligned} y(t) &= h(t) * Ue^{\mathrm{st}}1(t) \\ &= \begin{cases} H(s)Ue^{\mathrm{st}}1(t) + \dfrac{n_1}{d_0}U\delta(t) & \text{for } d_1 = 0 \\ H(s)Ue^{\mathrm{st}}1(t) - \dfrac{\left[d_1 n_0 - d_0 n_1\right]}{d_1 s + d_0}Ue^{-\frac{d_0}{d_1}}1(t) & \text{for } d_1 \neq 0 \end{cases} \end{aligned} \tag{9.28}$$

For $t > 0$, the SSS remains present, while there is another term of importance when $d_1 \neq 0$. This is a transient term, which disappears after a sufficient waiting time in the case of an asymptotically stable circuit. That is, the SSS is truly a steady state, although one may have to wait for it to dominate. If a nonzero zero-input response exists, it must be added to the right side of (9.28), but for $t > 0$ this is of the same form as the transient already present, therefore, the conclusion is identical (the SSS eventually predominates over the transient terms for an asymptotically stable circuit).

Because a cosine is the real part of a complex exponential and the real part is obtained as the sum of two terms, we can use linearity of the circuit to quickly obtain the output to a cosine input when we know the output due to an exponential. We merely write the input as the sum of two complex conjugate exponentials and then take the complex conjugates of the outputs that are summed. In the case of real coefficients in the transfer function, this is equivalent to taking the real part of the output when we take the real part of the input; that is, $y = \Re(h * u_3) = h * u$, when $u = \Re(u_e)$, if y is real for all real u.

Network Time Constant

The time constant, t_c, was defined earlier as the time for which a transient decays to $1/e$ of the intial value. As such, the time constant shows up in signals throughout the circuit and is a very useful parameter when identifying a circuit from its responses. In an RC circuit, the time constant physically results from the interaction of the equivalent capacitor (of which only one exists in a first-order circuit) of capacitance C_{eq}, and the Thévenin's equivalent resistor, of resistance R_{eq}, that it sees. Thus, $t_c = R_{eq}C_{eq}$.

Closely related to the time constant is the **rise time**. Considering the low-pass case, the rise time, t_r is defined as the time for the unit step response to go between 10% and 90% of its final value from its initial value. This is easily calculated because the unit step response is given by

$$y_{1(\cdot)}(t) = h(t) * 1(t) = \frac{n_0}{d_0}\left[1 - e^{-\frac{d_0}{d_1}t}\right] \cdot 1(t) \tag{9.29}$$

Assuming a stable circuit and setting this equal to 0.1 and 0.9 times the final value, n_0/d_0, it is readily found that

$$t_r = \frac{d_1}{d_0} \cdot \ln(9) = \left[\ln(9)\right] \cdot t_c \approx 2.2t_c \tag{9.30}$$

At this point, it is worth noting that for theoretical studies the time constant can be normalized to 1 by normalizing the time scale. Thus, assuming d_1 and $d_0 \neq 0$ the differential equation can be written as

$$d_0 \cdot \left[\frac{d_1}{d_0} \cdot \frac{dy}{d(d_1/d_0)\left(t(d_1/d_0)\right)} + y\right] = d_0\left[\frac{dy}{dt_n} + y\right] \tag{9.31}$$

where $t_n = (d_0/d_1)t$ is the normalized time.

References

[1] L. P. Huelsman, *Basic Circuit Theory with Digital Computations*, Englewood Cliffs, NJ: Prentice Hall, 1972.

[2] R. W. Newcomb and B. Dziurla, "Some circuits and systems applications of semistate theory," *Circuits, Systems, and Signal Processing*, vol. 8, no. 3, pp. 235–260, 1989.

9.3 Second-Order Circuits

Introduction

Because real transfer functions can be factored into real second-order transfer functions, second-order circuits are probably the most important circuits available; most designs are based upon them. As with first-order circuits, this chapter is limited to single-input-single-output linear time-invariant circuits, and unless otherwise stated, here real-valued quantities are assumed. By definition a **second-order circuit** is described by the differential equation

$$d_2 \cdot \frac{d^2y}{dt^2} + d_1 \cdot \frac{dy}{dt} + d_0 \cdot y = n_2 \cdot \frac{d^2u}{dt^2} + n_1 \cdot \frac{du}{dt} + n_0 \cdot u \tag{9.32}$$

where d_i and n_i are "denominator" and "numerator" constants, $i = 0, 1, 2$, which, unless mentioned to the contrary, are taken to be real. Continuing the notation used for first-order circuits, $y = y(\cdot)$ is the output and $u = u(\cdot)$ is the input; both u and y are generalized functions of time t. Assume that $d_2 \neq 0$,

which is the normal case because any of the other special cases can be considered as cascades of real degree one circuits.

Again, treating the derivative as an operator, $p = d[\cdot]/dt$, (9.32) is written as

$$y = \frac{n_2 p^2 + n_1 p + n_0}{d_2 p^2 + d_1 p + d_0} u \tag{9.33}$$

with the **transfer function**

$$\begin{aligned} H(p) &= \frac{1}{d_2}\left[\frac{n_2 p^2 + n_1 p + n_0}{p^2 + (d_1/d_2)p + (d_0/d_2)}\right] \\ &= \frac{1}{d_2}\left[n_2 + \frac{(n_1 - (d_1/d_2)n_2)p + (n_0 - (d_0/d_2)n_2)}{p^2 + (d_1/d_2)p + (d_0/d_2)}\right] \end{aligned} \tag{9.34}$$

where the second form results by long division of the denominator into the numerator. Because they occur most frequently when second-order circuits are discussed, we rewrite the denominator in two equivalent customarily used forms:

$$p^2 + \frac{d_1}{d_2}p + \frac{d_0}{d_2} = p^2 + \frac{\omega_n}{Q}p + \omega_n^2 = p^2 + 2\zeta\omega_n p + \omega_n^2 \tag{9.35}$$

where ω_n is the undamped natural frequency ≥ 0, Q is the quality factor, and ζ is the damping factor $= 1/(2Q)$. The transfer function is accordingly

$$H(p) = \frac{1}{d_2}\left[\frac{n_2 p^2 + n_1 p + n_0}{p^2 + (\omega_n/Q)p + \omega_n^2}\right] = \frac{1}{d_2}\left[\frac{n_2 p^2 + n_1 p + n_0}{p^2 + 2\zeta\omega_n p + \omega_n^2}\right] \tag{9.36}$$

Table 9.2 lists several of the more important transfer functions, which, as in the first-order case, are operators as functions of the derivative operator p.

Zero-Input and Zero-State Response

Again, as in the first-order case, a convenient tool for investigating the time-domain behavior of a second-order circuit is the **state variable description**. Letting the state vector be $x(\cdot)$, the state-space representation is

$$\begin{aligned} px &= Ax + Bu \\ y &= Cx + Du \end{aligned} \tag{9.37}$$

where, as above, $p = d[\cdot]/dt$, and A, B, C, D are constant matrices. In the present case, these matrices are real and one convenient choice, among many, is

$$\begin{aligned} px &= \begin{bmatrix} 0 & 1 \\ -\dfrac{d_0}{d_2} & -\dfrac{d_1}{d_2} \end{bmatrix} x + \begin{bmatrix} n_1 - \dfrac{d_1}{d_2}n_2 \\ \left(n_0 - \dfrac{d_0}{d_2}n_2\right) - \left(n_1 - \dfrac{d_1}{d_2}n_2\right) \end{bmatrix} u \\ y &= \begin{bmatrix} \dfrac{1}{d_2} & 0 \end{bmatrix} x + \begin{bmatrix} \dfrac{n_1}{d_2} \end{bmatrix} u \end{aligned} \tag{9.38}$$

TABLE 9.2 Typical Second-Order Circuit Transfer Functions

Transfer Function	Description	Impulse Response
$\frac{n_0}{d_2}\frac{1}{p^2+2\zeta\omega_n p+\omega_n^2}$	Low-pass	$h_{lp}(t)=\frac{n_0}{d_2}\frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}\omega_n}\sin\left(\sqrt{1-\zeta^2}\omega_n t\right)1(t)$
	High-pass	
$\frac{n_2}{d_2}\frac{p^2}{p^2+2\zeta\omega_n p+\omega_n^2}$	$\theta=\arctan 2\left(\frac{\zeta}{\sqrt{1-\zeta^2}}\right)$	$h_{hp}(t)=\frac{n_2}{d_2}\left[\delta(t)-\frac{\omega_n e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\sin\left(\sqrt{1-\zeta^2}\omega_n t+2\theta\right)1(t)\right]$
	Bandpass	
$\frac{n_1}{d_2}\frac{p}{p^2+2\zeta\omega_n p+\omega_n^2}$	$\theta=\arctan 2\left(\frac{\zeta}{\sqrt{1-\zeta^2}}\right)$	$h_{bp}(t)=\frac{n_1}{d_2}\frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\cos\left(\sqrt{1-\zeta^2}\omega_n t+\theta\right)1(t)$
$\frac{n_2}{d_2}\frac{p^2+\omega_0^2}{p^2+2\zeta\omega_n p+\omega_n^2}$	Band-stop	$h_{bs}(t)=h_{hp}(t)+\frac{n_2\omega_0^2}{n_0}h_{lp}(t)$
$\frac{n_2}{d_2}\frac{p^2-2\zeta\omega_n p+\omega_n^2}{p^2+2\zeta\omega_n p+\omega_n^2}$	All-pass	$h_{ap}(t)=\frac{n_2}{d_2}\left[\delta(t)-\frac{4\zeta\omega_n e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}}\cos\left(\sqrt{1-\zeta^2}\omega_n t+\theta\right)1(t)\right]$
$\frac{n_0}{d_2}\frac{1}{p^2+\omega_n^2}$	Oscillator,	$h_{osc}(t)=\frac{n_0}{d_2}\sin(\omega_n t)\cdot 1(t)$
	when u = 0	$y(t)\big\|_{u=0}=y(0)\cdot\cos(\omega_n t)+\frac{y'(0)}{\omega_n}\cdot\sin(\omega_n t)$

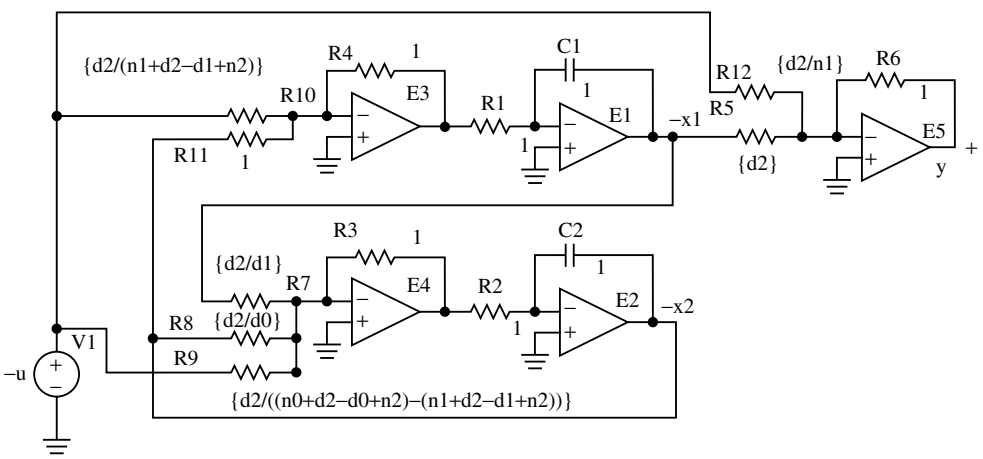


FIGURE 9.2 Generic, second-order op-amp RC circuit.

Here, the state is the 2-vector $x = [x_1\ x_2]^T$, with the superscript T denoting transpose. Normally, the state would consist of capacitor voltages and/or inductor currents, although at times one may wish to use linear combinations of these. From these state variable equations, a generic operational-amplifier (op-amp) RC circuit to realize any of this class of second-order circuits is readily designed and given in Figure 9.2. In the figure, all voltages are referenced to ground and normalized capacitor and resistor values are listed. Alternate designs in terms of only CMOS differential pairs and capacitors can also be given [3], while a number of alternate circuits exist in the catalog of Sallen and Key [4].

Because (9.38) represents a set of linear constant coefficient differential equations, superposition applies and its solution can again be broken into two parts, the part due to initial conditions, $x(0)$, called the zero-input response, and the part due solely to the input u, the zero-state response.

The **zero-input response** is readily found by solving the state equations with $u = 0$ and initial conditions $x(0)$. The result is $y(t) = C\exp(At)\,x(0)$, which can be evaluated by several means, including the following. Using a prime to designate the time derivative, first note that when $u = 0$, $x_1(t) = d_2\,y(t)$ and (from the first row of A) $x_1(t)' = x_2(t) = d_2\,y(t)'$. Thus, $x_1(0) = d_2\,y(0)$ and $x_2(0) = d_2\,y'(0)$, which allow the initial conditions to be expressed in terms of the measurable output quantities. To evaluate $\exp(At)$, note that its terms are linear combinations of terms with complex frequencies that are zeroes of the characteristic polynomial

$$\begin{aligned}\det(s1_2 - A) &= \det\begin{pmatrix} s & -1 \\ \omega_n^2 & s + 2\zeta\omega_n \end{pmatrix} = s^2 + 2\zeta\omega_n s + \omega_n^2 \\ &= (s - s_-)(s - s_+)\end{aligned} \tag{9.39}$$

For which the roots, called **natural frequencies**, are

$$s_\pm = \left(-\zeta \pm \sqrt{\zeta^2 - 1}\right)\omega_n = \left(-1 \pm \sqrt{1 - 4Q^2}\right)\frac{\omega_n}{2Q} \tag{9.40}$$

The case of equal roots will only occur when $\zeta^2 = 1$, which is the same as $Q^2 = 1/4$, for which the roots are real. Indeed, if the damping factor, ζ, is > 1 in magnitude, or equivalently, if the quality factor, Q, is $<1/2$ in magnitude, the roots are real and the circuit can be considered a cascade of two first-order circuits. Thus, assume here and in the following that unless otherwise stated, $Q^2 > 0.25$, which is the same as $\zeta^2 < 1$, in which case the roots are complex conjugates, $s_- = s_+^*$

$$s_\pm = \left(-\zeta \pm j\sqrt{1 - \zeta^2}\right)\omega_n = \left(-1 \pm j\sqrt{4Q^2 - 1}\right)\frac{\omega_n}{2Q}, \quad j = \sqrt{-1} \tag{9.41}$$

By writing $y(t) = a \cdot \exp(s_+ t) + b \cdot \exp(s_- t)$, for unknown constants a and b, differentiating and setting $t = 0$ we can solve for a and b, and after some algebra and trigonometry obtain the zero-input response

$$y(t) = \frac{e^{-\zeta\omega_n t}}{\sqrt{1 - \zeta^2}}\left\{y(0)\cdot\cos\left(\sqrt{1 - \zeta^2}\,\omega_n t - \theta\right) + \frac{y'(0)}{\omega_n}\cdot\sin\left(\sqrt{1 - \zeta^2}\,\omega_n t\right)\right\} \tag{9.42}$$

where $\theta = \arctan2(\zeta/\sqrt{1 - \zeta^2})$with arctan2(·) being the arc tangent function that incorporates the sign of its argument.

The form given in (9.42) allows for some useful observations. Remembering that this assumes $\zeta^2 < 1$, first note that if no damping occurs, that is, $\zeta = 0$, then the natural frequencies are purely imaginary, $s_+ = j\omega_n$ and $s_- = -s_+$, and the response is purely oscillatory, taking the form shown in the last line of Table 9.2. If the damping is positive, as it would be for a passive circuit having some loss, usually via positive resistors, then the natural frequencies lie in the left half s-plane, and y decays to zero at positive infinite time so that any transients in the circuit die out after a sufficient wait. The circuit is then called **asymptotically stable**. However, if the damping is negative, as it could be for some positive feedback circuits or those with negative resistance, then the response to nonzero initial conditions increases in amplitude without bound, although in an oscillatory manner, as time increases, and the circuit is said to be **unstable**. In the unstable case, as time decreases through negative time the amplitude also damps out to zero, but usually the responses backward in time are not of as much interest as those forward in time.

For the **zero-state response**, the impulse response, $h(t)$, is convoluted with the input, that is, $y = h * u$, for which we can use the fact that $h(t)$ is the inverse Laplace transform of $H(s) = C[sI_2 - A]^{-1}B$. The denominator of $H(s)$ is $\det(sI_2 - A) = s^2 + 2\zeta\omega_n s + \omega_n^2$, for which the causal inverse Laplace transform is

$$e^{s_+ t}1(t) * e^{s_- t}1(t) = \begin{cases} \dfrac{e^{s_+ t} - e^{s_- t}}{s_+ - s_-}1(t) & \text{if } s_- \neq s_+ \\ te^{s_+ t}1(t) & \text{if } s_- = s_+ \end{cases} \tag{9.43}$$

Here, the bottom case is ruled out when only complex natural frequencies are considered, following the assumption of handling real natural frequencies in first-order circuits, made previously. Consequently,

$$e^{s_+ t}1(t) * e^{s_- t}1(t) = \frac{e^{s_+ t} - e^{s_- t}}{s_+ - s_-}1(t) = \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}\omega_n}\sin\left(\sqrt{1-\zeta^2}\omega_n t\right)\cdot 1(t) \tag{9.44}$$

Again, assuming $\zeta^2 < 1$ using the preceding calculations give the zero-state response as

$$\begin{aligned} y(t) &= \frac{1}{d_2}\left\{\frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}\omega_n}\sin\left(\sqrt{1-\zeta^2}\omega_n t\right)1(t) * \right. \\ &\qquad \left.\left[\left(n_1 - \frac{d_1}{d_2}n_2\right)\delta'(t) + \left(n_0 - \frac{d_0}{d_2}n_2\right)\delta(t)\right] + n_2\delta(t)\right\} * u(t) \\ &= \frac{1}{d_2}\left\{\frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}\omega_n}\sin\left(\sqrt{1-\zeta^2}\omega_n t\right)1(t) * \right. \\ &\qquad \left.\left[n_2\delta''(t) + n_1\delta'(t) + n_0\delta(t)\right]\right\} * u(t) \end{aligned} \tag{9.45}$$

The bottom equivalent form is easily seen to result from writing the transfer function $H(p)$ as the product of two terms $1/[d_2(p^2 + 2\zeta\omega_n p + \omega_n^2)$ and $[n_2 p^2 + n_1 p + n_0]$ convoluting the causal impulse response (the inverse of the left half-plane converging Laplace transform), of each term. From (9.45), we directly read the **impulse response** to be

$$\begin{aligned} h(t) = \frac{1}{d_2}\left\{\frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}\omega_n}\sin\left(\sqrt{1-\zeta^2}\omega_n t\right)1(t)\right. \\ \left. * \left[n_2\delta''(t) + n_1\delta'(t) + n_0\delta(t)\right]\right\} \end{aligned} \tag{9.46}$$

Equations (9.45) and (9.46) are readily evaluated further by noting that the convolution of a function with the second derivative of the impulse, the first derivative of the impulse, and the impulse itself is the second derivative of the function, the first derivative of the function, and the function itself, respectively. For example, in the low-pass case we find the impulse response to be, using (9.46),

$$h_{lp}(t) = \frac{n_0}{d_2}\frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}\omega_n}\sin\left(\sqrt{1-\zeta^2}\omega_n t\right)1(t) \tag{9.47}$$

By differentiating we find the bandpass and then high-pass impulse responses to be, respectively,

$$h_{bp}(t) = \frac{n_1}{d_2} \frac{e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}} \cos\left(\sqrt{1-\zeta^2}\,\omega_n t + \theta\right) 1(t) \tag{9.48}$$

$$h_{hp}(t) = \frac{n_2}{d_2}\left[\delta(t) - \frac{\omega_n e^{-\zeta\omega_n t}}{\sqrt{1-\zeta^2}} \sin\left(\sqrt{1-\zeta^2}\,\omega_n t + 2\theta\right) 1(t)\right] \tag{9.49}$$

In both cases, the added phase angle is given, as in the zero input response, via $\theta = \arctan2(\zeta/\sqrt{1-\zeta^2})$. By adding these last three impulse responses suitably scaled the impulse responses of the more general second-order circuits are obtained.

Some comments on **normalizations** are worth mentioning in passing. Because $d_2 \neq 0$, one could assume d_2 to be 1 by absorbing its actual value in the transfer function numerator coefficients. If $\omega_n \neq 0$, time could also be scaled so that $\omega_n = 1$ could be taken, in which case a normalized time, t_n, is introduced. Thus, $t = \omega_n t_n$ and, along with normalized time, comes a normalized differential operator $p_n = d[\cdot]/dt_n = d[\cdot]/d(t/\omega_n) = \omega_n p$. This, in turn, leads to a normalized transfer function by substituting $p = p_n/\omega_n$ into $H(p)$. Thus, much of the treatment could be carried out on the normalized transfer function x

$$H_n(p_n) = H(p) = \frac{n_{2n}p_n^2 + n_{1n}p_n + n_{0n}}{p_n^2 + 2\zeta p_n + 1} \qquad p_n = \omega_n p \tag{9.50}$$

In this normalized form, it appears that the most important parameter in fixing the form of the response is the damping factor $\zeta = 1/(2Q)$.

Transient and Steady-State Responses

Let us now excite the circuit with an eternal exponential input, $u(t) = U\exp(st)$ for $-\infty < t < \infty$ at the complex frequency $s = \sigma + j\omega$, where s is chosen as different from either of the natural frequencies, $s_{\pm}$, and U is a constant. As with the first-order and, indeed, any higher-order, case the response is $y(t) = Y(s)\exp(st)$, as is observed by direct substitution into (9.32). This substitution yields directly

$$Y(s) = \frac{1}{d_2}\left[\frac{n_2 s^2 + n_1 s + n_0}{s^2 + 2\zeta\omega_n s + \omega_n^2}\right] \cdot U \tag{9.51}$$

where $y(t) = Y(s)\exp(st)$ for $u(t) = U\exp(st)$ over $-\infty < t < \infty$. That is, an exponential excitation yields an exponential response at the same (complex) frequency $s = \sigma + j\omega$ as that for the input, as long as s is not one of the two natural frequencies. (s may have positive as well as negative real parts and is best considered as a frequency and not as the Laplace transform variable because the latter is limited to regions of convergence.) Because the denominator polynomial of $Y(s)$ has roots which are the natural frequencies, the magnitude of Y becomes infinite as the frequency of the excitation approaches s_+ or s_-. Thus, the natural frequencies s_+ and s_- are also called **poles** of the transfer function.

When $\sigma = 0$ the excitation and response are both sinusoidal and the resulting response is called the **sinusoidal steady state** (SSS). From (9.51), the SSS response is found by substituting the complex frequency $s = j\omega$ into the transfer function, now evaluated on complex numbers rather than differential operators as above,

$$H(s) = \frac{1}{d_2}\left[\frac{n_2 s^2 + n_1 s + n_0}{s^2 + 2\zeta\omega_n s + \omega_n^2}\right] \tag{9.52}$$

Next, an exponential input is applied, which starts at $t = 0$ instead of at $t = -\infty$; i.e., $u(t) = U\exp(st)1(t)$. Then, the output is found by using the convolution $y = h * u$, which, from the discussion at (9.45), is expressed as

$$y(t) = h * u = \frac{1}{d_2} e^{s_+ t} 1(t) * e^{s_- t} 1(t) * \left[n_2 \delta''(t) + n_1 \delta'(t) + n_0 \delta(t) \right] * e^{st} 1(t)$$

$$= H(s) U e^{st} 1(t) + \left\{ \frac{1}{d_2 (s_+ - s_-)} \left[\left(\frac{N(s)}{s_+ - s} + n_2 (s + s_+) + n_1 \right) e^{s_+ t} \right. \right. \tag{9.53}$$

$$\left. \left. - \left(\frac{N(s)}{s_+ - s} + n_2 (s + s_-) + n_1 \right) e^{s_- t} \right] 1(t) \right\}$$

in which $N(s)$ is the numerator of the transfer function and we have assumed that s is not equal to a natural frequency. The second term on the right within the braces varies at the natural frequencies and as such is called the **transient response**, while the first term is the term resulting directly from an eternal exponential, but now with the negative time portion of the response removed. If the system is stable, the transient response decays to zero as time increases and, thus, if we wait long enough the transient response of a stable system can be ignored if the complex frequency of the input exponential has a real part that is greater than that of the natural frequencies. Such is the case for exponentials that yield sinusoids; in that case $\sigma = 0$, or $s = j\omega$. In other words, for an asymptotically stable circuit the output approaches that of the SSS when the input frequency is purely imaginary. If we were to excite at a natural frequency then the first part of (9.53) still could be evaluated using the time-multiplied exponential of (9.43); however, the transient and the steady state are now mixed, both being at the same "frequency."

Because actual sinusoidal signals are real, we use superposition and the fact that the real part of a complex signal is given by adding complex conjugate terms:

$$\cos(\omega t) = \Re\left[e^{j\omega t} \right] = \frac{e^{j\omega t} + e^{-j\omega t}}{2} \tag{9.54}$$

This leads to the SSS response for an asymptotically stable circuit excited by $u(t) = U \cos(\omega t) 1(t)$ to be

$$\begin{aligned} y(t) &= \frac{H(j\omega) U e^{j\omega t} + H(-j\omega) U^* e^{-j\omega t}}{2} \\ &= |H(j\omega)||U| \cos\left(\omega t + \angle H(j\omega) + \angle U \right) \end{aligned} \tag{9.55}$$

Here, we assumed that the circuit has real-valued components such that $H(-j\omega)$ is the complex conjugate of $H(j\omega)$. In which case, the second term in the middle expression is the complex conjugate of the first.

Network Characterization

Although the impulse response is useful for theoretical studies, it is difficult to observe it experimentally due to the impossibility of creating an impulse. However, the unit step response is readily measured, and from it the impulse response actually can be obtained by numerical differentiation if needed. However, it is more convenient to work directly with the unit step response and, consequently, practical characterizations can be based upon it. The treatment most conveniently proceeds from the normalized low-pass transfer function

$$H(p) = \frac{1}{p^2 + 2\zeta p + 1}, \quad 0 < \zeta < 1 \tag{9.56}$$

The **unit step response** follows by applying the input $u(t) = 1(t)$ and noting that the unit step is the special case of an exponential multiplied unit step, where the frequency of the exponential is zero. Conveniently, (9.43) can be used to obtain

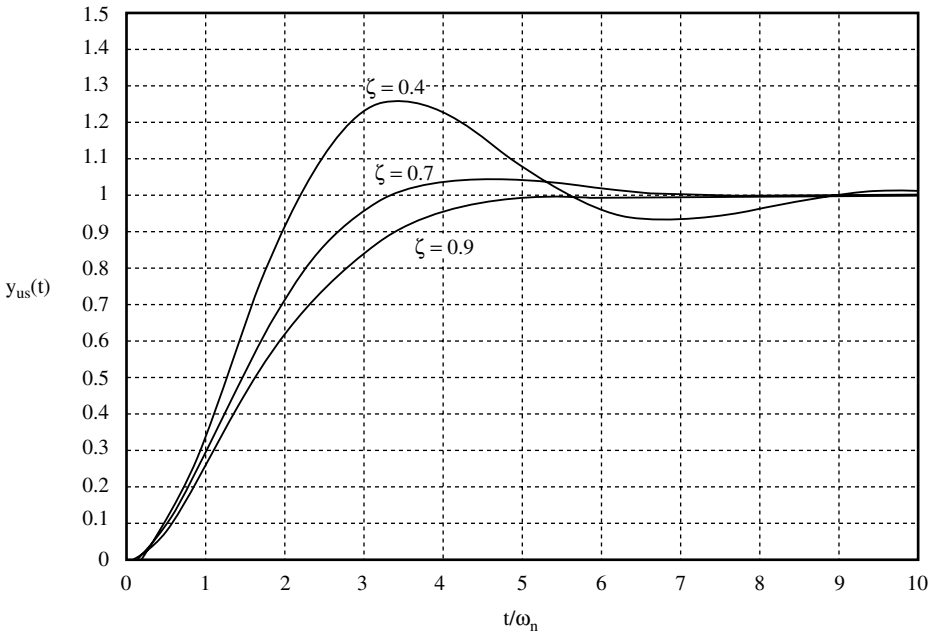


FIGURE 9.3 Unit step response for different damping factors.

$$y_{us}(t) = 1(t) - \frac{e^{-\zeta t}}{\sqrt{1-\zeta^2}} \cos\left(\sqrt{1-\zeta^2}\, t - \theta\right) \cdot 1(t), \quad \theta = \arctan\left(\frac{\zeta}{\sqrt{1-\zeta^2}}\right) \tag{9.57}$$

Typical unit step responses are plotted in Figure 9.3 where, for a small damping factor, overshoot can be considerable, with oscillations around the final value and in addition, a long settling time before reaching the final value. In contrast, with a large damping factor, although no overshoot or oscillation occurs, the rise to the final value is long. A compromise for obtaining a quick rise to the final value with no oscillations is given by choosing a damping factor of 0.7, this being called the **critical value**; i.e., **critical damping** is $\zeta_{crit} = 0.7$.

References

[1] L. P. Huelsman, *Basic Circuit Theory with Digital Computations*, Englewood Cliffs, NJ: Prentice Hall, 1972.

[2] V. I. Arnold, *Ordinary Differential Equations*, Cambridge, MA: MIT Press, 1983.

[3] J. E. Kardontchik, *Introduction to the Design of Transconductor-Capacitor Filters*, Boston: Kluwer Academic Publishers, 1992.

[4] R. P. Sallen and E. L. Key, "A practical method of designing RC active filters," *IRE Trans. Circuit Theory*, vol. CT-2, no. 1, pp. 74–85, March 1955.

10

State-Variable Techniques

K. S. Chao
Texas Tech University

10.1 The Concept of States

For resistive (or memoryless) circuits, given the circuit structure, the present output depends only on the present input. In order to analyze a dynamic circuit, however, in addition to the present input it is also necessary to know the state of the circuit at some time t_0. The state of the circuit at t_0 represents the condition of the circuit at $t = t_0$, and is related to the energy storage of the circuit, or the voltage (or electric charge) across the capacitor and the currents (or magnetic fluxes) through the inductors. These voltages and currents are considered as the state of the circuit at $t = t_0$. For $t > t_0$, the behavior of the circuit is completely characterized by these variables. In view of the preceding, a definition for the state of a circuit can now be given.

Definition: The state of a circuit at time t_0 is the minimum amount of information at t_0 that, along with the input to the circuit for $t \geq t_0$, uniquely determines the behavior of the circuit for $t \geq t_0$.

The concept of states is closely related to the order of complexity of the circuit. The order of complexity of a circuit is the minimum number of initial conditions which, along with the input, is sufficient to determine the future behavior of the circuit. Furthermore, if a circuit is described by an n^{th}-order linear differential equation, it is well known that the general solution for $t \geq t_0$ contains n arbitrary constants which are determined by n initial conditions. This set of n initial conditions contains information concerning the circuit prior to $t = t_0$ and constitutes the state of the circuit at $t = t_0$. Thus, the order of complexity or the order of a circuit is the same as the order of the differential equation that describes the circuit, and it is also the same as the number of state variables that can be defined in a circuit. For an n^{th}-order circuit, the state of the circuit at $t = t_0$ consists of a set of n numbers that denotes a vector in an n-dimensional state space spanned by the n corresponding state variables. This key number n can simply be obtained by inspection of the circuit. Knowing the total number of energy storage elements, n_{LC}, the total number of independent capacitive loops, n_C, and the total number of independent inductive cutsets, n_L, the order of complexity n of a circuit is given by

$$n = n_{LC} - n_L - n_C \tag{10.1}$$

A capacitive loop is defined as one that consists of only capacitors and possibly voltage sources while an inductive cutset represents a cutset that contains only inductors and possibly current sources. The following two examples illustrate the concept of states.

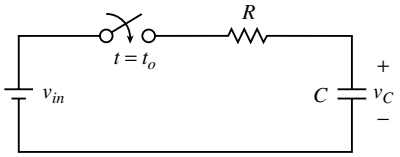


FIGURE 10.1 A simple RC circuit.

Example 1. Consider a simple RC-circuit in Figure 10.1. The circuit equation is

$$RC\,\frac{dv_c(t)}{dt}+v_c(t)=v_{in} \qquad \text{for} \qquad t\geq t_0 \tag{10.2}$$

and the corresponding capacitor voltage is easily obtained as

$$v_c(t)=\left[v_c(t_0)-v_{in}\right]e^{-\frac{1}{RC}(t-t_0)}+v_{in} \quad \text{for} \quad t\geq t_0 \tag{10.3}$$

For this first-order circuit, it is clear from (10.3) the capacitor voltage for $t \geq t_0$ is uniquely determined by the initial condition $v_c(t_0)$ and the input voltage v_{in} for $t \geq t_0$. This is independent of the charging circuit for the capacitor prior to t_0. Hence, $v_c(t_0)$ is the state of the circuit at $t = t_0$ and $v_c(t)$ is regarded as the state variable of the circuit.

Example 2. As another illustration, consider the circuit of Figure 10.2, which is a slight modification of the circuit considered in the previous example. The circuit equation and its corresponding solution are readily obtained as

$$\frac{dv_{C_1}(t)}{dt}=-\frac{1}{R(C_1+C_2)}v_{C_1}(t)+\frac{1}{R(C_1+C_2)}v_{in} \tag{10.4}$$

and

$$v_{C_1}(t)=\left(v_{C_1}(t_0)-v_{in}\right)e^{\frac{-1}{R(C_1+C_2)}(t-t_0)}+v_{in} \quad \text{for} \quad t\geq t_0 \tag{10.5}$$

respectively. Even though two energy storage elements exist, one can only arbitrarily specify one independent initial condition. Once the initial condition on C_1, $v_{C_1}(t_0)$, is specified, the initial voltage on C_2 is automatically constrained by the loop equation $v_{C_2}(t) = V_{C_1}(t) - E$ at t_0. The circuit is thus still first order and only one state variable can be assigned for the circuit. It is clear from (10.5) that with the input v_{in}, $v_{C_1}(t_0)$ is the minimum amount of information that is needed to uniquely determine the behavior of this circuit. Hence, $v_{C_1}(t)$ is the state variable of the circuit. One can just as well analyze the circuit by solving a first-order differential equation in terms of $v_{C_2}(t)$ with $v_{C_2}(t_0)$ defined as the state of the circuit at $t = t_0$. The selection of state variables is thus not unique. In this example, either $v_{C_1}(t)$ or $v_{C_2}(t)$ can be defined as the state variable of the circuit. In fact, it is easily shown that any linear combination of $v_{C_1}(t)$ and $v_{C_2}(t)$ can also be regarded as state variables.

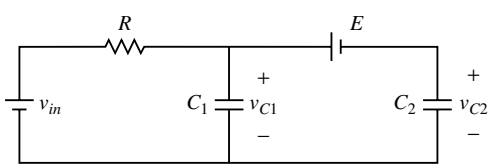


FIGURE 10.2 The circuit for Example 2.

10.2 State-Variable Formulation via Network Topology

Various mathematical descriptions of circuits are available. Depending on the type of analysis used, different formulations of circuit equations may result. In the state variable formulation, a system of n first-order differential equations is written in the form

$$\dot{\mathbf{x}} = \mathbf{f}(\mathbf{x}, t) \tag{10.6}$$

where $\mathbf{x}$ is an n x 1 vector consisting of n state variables for an n^{th}-order circuit and t represents the time variable. This set of equations is usually referred to as the state equation in normal form.

When compared with other circuit descriptions, the state-variable representation is not necessarily the simplest. It does, however, simultaneously provide the solution of all state variables and hence yields the behavior of the entire circuit. The state equation is also particularly suitable for analysis by numerical techniques. Another distinct advantage of the state-variable approach is that it can be easily extended to nonlinear and/or time varying circuits.

Example 3. Consider the linear circuit of Figure 10.3. By inspection, the order of complexity of this circuit is three. Hence, three state variables are selected as $x_1 = v_{C_1}$, $x_2 = v_{C_2}$, and $x_3 = i_L$. Because the left-hand side of the normal form equation is the derivative of the state vector, it is necessary to express the voltage across the inductors and the currents through the capacitors in terms of the state variables and the input sources.

The current through C_1 can be obtained by writing a Kirchhoff's current law (KCL) equation at node 1 to yield

$$\begin{aligned} C_1 \frac{dv_{C_1}}{dt} &= i_{R_1} - i_L \\ &= \frac{1}{R_1}\left(v_s - v_{C_1}\right) - i_L \end{aligned}$$

or

$$\frac{dv_{C_1}(t)}{dt} = -\frac{1}{R_1 C_1} v_{C_1} - \frac{1}{C_1} i_L + \frac{1}{R_1 C_1} v_s \tag{10.7}$$

In a similar manner, applying KCL to node 2 gives

$$\begin{aligned} C_2 \frac{dv_{C_2}}{dt} &= i_L - i_{R_3} + i_s \\ &= i_L - \frac{1}{R_3} v_{C_2} + i_s \end{aligned}$$

or

$$\frac{dv_{C_2}(t)}{dt} = -\frac{1}{R_3 C_2} v_{C_2} + \frac{1}{C_2} i_L + \frac{1}{C_2} i_s \tag{10.8}$$

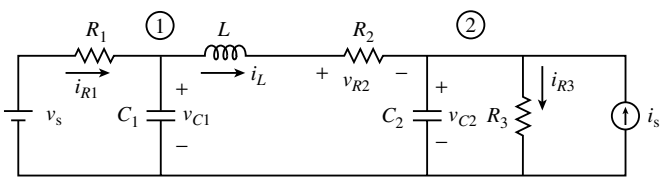


FIGURE 10.3 The circuit for Example 3.

The expression for the inductor voltage is derived by applying KVL to the mesh containing L, R_2, C_2, and C_1 yielding

$$L\frac{di_L}{dt} = v_{C_1} - v_{C_2} - R_2 i_L$$

or

$$\frac{di_L}{dt} = \frac{1}{L}v_{C_1} - \frac{1}{L}v_{C_2} - \frac{R_2}{L}i_L \tag{10.9}$$

Equations (10.7), (10.8), and (10.9) are the state equations that can be expressed in matrix form as

$$\begin{bmatrix} \frac{dv_{C_1}}{dt} \\ \frac{dv_{C_2}}{dt} \\ \frac{di_L}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{1}{R_1C_1} & 0 & -\frac{1}{C_1} \\ 0 & -\frac{1}{R_3C_2} & \frac{1}{C_2} \\ \frac{1}{L} & -\frac{1}{L} & -\frac{R_2}{L} \end{bmatrix} \begin{bmatrix} v_{c_1} \\ v_{c_2} \\ i_L \end{bmatrix} + \begin{bmatrix} \frac{1}{R_1C_1} & 0 \\ 0 & \frac{1}{C_2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_s \\ i_s \end{bmatrix} \tag{10.10}$$

Any number of branch voltages and/or currents may be chosen as output variables. If i_{R1} and v_{R2} are considered as outputs for this example, then the output equations, written as a linear combination of state variables and input sources become

$$i_{R_1} = \frac{1}{R_1}\left(v_s - v_{C_1}\right) \tag{10.11}$$

$$v_{R_2} = R_2 i_L \tag{10.12}$$

or in matrix form

$$\begin{bmatrix} i_{R_1} \\ v_{R_2} \end{bmatrix} = \begin{bmatrix} -\frac{1}{R_1} & 0 & 0 \\ 0 & 0 & R_2 \end{bmatrix} \begin{bmatrix} v_{C_1} \\ v_{C_2} \\ i_L \end{bmatrix} + \begin{bmatrix} \frac{1}{R_1} & 0 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_s \\ i_s \end{bmatrix} \tag{10.13}$$

In general, for an n^{th}-order linear circuit with r input sources and m outputs, the state and output equations are represented by

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}\mathbf{u} \tag{10.14}$$

and

$$\mathbf{y} = \mathbf{C}\mathbf{x} + \mathbf{D}\mathbf{u} \tag{10.15}$$

where $\mathbf{x}$ is an $n \times 1$ state vector, $\boldsymbol{u}$ is an $r \times 1$ vector representing the r input sources, $m \times 1$-vector $\mathbf{y}$ denotes the m output variables, $\mathbf{A}$, $\mathbf{B}$, $\mathbf{C}$, and $\mathbf{D}$ are of order $n \times n$, $n \times r$, $m \times n$, and $m \times r$, respectively.

In the preceding example, the state equations are obtained by inspection for a simple circuit by writing voltage equations for inductors and current equations for capacitors and properly eliminating the nonstate variables. For more complicated circuits, a systematic procedure for eliminating the nonstate variables is desirable. Such a procedure can be generated with the aid of a proper tree. A proper tree is a tree obtained from the associated network graph that contains all capacitors, independent voltage sources, and possibly some resistive elements, but does not contain inductors and independent current sources.

The selection of such a tree is always possible if the circuit contains no capacitive loops and no inductive cutsets. The reason for providing such a tree for writing state equations is obvious. With each tree branch, there is a unique cutset known as the fundamental cutset that contains only one tree branch and some links. Thus, if capacitors are in the tree, a fundamental cutset equation may be written for the corresponding currents through the capacitors. Similarly, every link (together with some tree branches) forms a unique loop called a fundamental loop. If inductors are selected as links, inductor voltages may be obtained by writing the corresponding fundamental loop equations. With the selection of a proper tree, state variables can be defined as the capacitor tree-branch voltages and inductive link currents. In view of the above observation, a systematic procedure for writing state equations can now be stated as follows:

STEP 1: From the associated directed graph, pick a proper tree.

STEP 2: Write fundamental cutset equations for the capacitive tree branches and express the capacitor currents in terms of link currents.

STEP 3: Write fundamental loop equations for the inductive links and express the inductor voltages in terms of tree-branch voltages.

STEP 4: Define the state variables. Capacitive tree-branch voltages and inductive link currents are selected as state variables. Other quantities such as capacitor charges and inductor fluxes may also be used.

STEP 5: Group the branch relations and the remaining fundamental equations according to their element types into three sets: resistor, inductor, and capacitor equations. Solve for the nonstate variables that appeared in the equations obtained in Steps 2 and 3 from the corresponding set of equations in terms of the state variables and independent sources.

STEP 6: Substitute the result of Step 5 into the equations obtained in Steps 2 and 3, and rearrange them in normal form.

Example 4. Consider again the same circuit in Figure 10.3. The various steps outlined previously are used to write the state equations.

STEP 1: The associated graph and the proper tree of the circuit are shown in Figure 10.4. The tree branches include v_s, C_1, C_2, and R_2.

STEP 2: The fundamental cutset associated with C_1 consists of tree branch C_1 and two links R_1 and L. By writing the current equation for this cutset, the capacitor current i_{c_1} is expressed in terms of link currents as

$$i_{C_1} = i_{R_1} - i_L \tag{10.16}$$

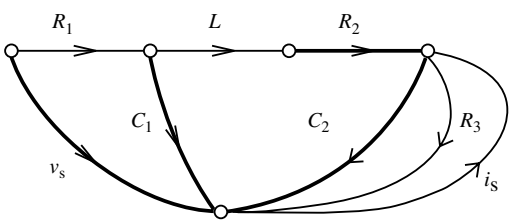


FIGURE 10.4 The directed graph associated with the circuit of Figure 10.3.

Similarly, the fundamental cutset $\{L, C_2, R_3, i_s\}$ associated with C_2 leads to

$$i_{C_2} = i_L - i_{R_3} + i_s \tag{10.17}$$

STEP 3: The fundamental loop associated with link L consists of L and tree branches R_2, C_2, and C_1. By writing the voltage equation around this loop, the inductor voltage can be written in terms of tree-branch voltages as

$$v_L = v_{C_1} - v_{C_2} - v_{R_2} \tag{10.18}$$

STEP 4: The tree-branch capacitor voltages v_{C_1}, V_{C_2},and inductive link current i_L are defined as the state variables of the circuit.

STEP 5: The branch relation and the remaining two fundamental loops for R_1 and R_2, and the fundamental cutset equation for R_2 are grouped into three sets.

Resistor equations:

$$v_{R_1} + v_{C_1} - v_s = 0 \tag{10.19}$$

$$i_{R_1} = \frac{1}{R_1} v_{R_1} \tag{10.20}$$

$$i_{R_2} - i_L = 0 \tag{10.21}$$

$$v_{R_2} = R_2 i_{R_2} \tag{10.22}$$

$$v_{R_3} - v_{C_2} = 0 \tag{10.23}$$

$$v_{R_3} = R_3 i_{R_3} \tag{10.24}$$

Inductor equations:

$$\phi_L = L i_L \quad \text{or} \quad v_L = \frac{d\phi_L}{dt} = L \frac{di_L}{dt} \tag{10.25}$$

Capacitor equations:

$$q_1 = C_1 v_{C_1} \quad \text{or} \quad i_{C_1} = \frac{dq_1}{dt} = C_1 \frac{dv_{C_1}}{dt} \tag{10.26}$$

$$q_2 = C_2 v_{C_2} \quad \text{or} \quad i_{C_2} = \frac{dq_2}{dt} = C_2 \frac{dv_{C_2}}{dt} \tag{10.27}$$

The resistive link currents r_{R_1}, i_{R_3}, and resistive tree-branch voltage V_{R_2} are solved from (10.19)–(10.24) in terms of the inductive link current i_L, the capacitive tree-branch voltages v_{C_1} and v_{C_2}, and sources as

$$i_{R_1} = \frac{1}{R_1}\left(v_s - v_{C_1}\right) \tag{10.28}$$

$$i_{R_3} = \frac{1}{R_3} v_{C_2} \tag{10.29}$$

and

$$v_{R_2} = R_2 i_L \tag{10.30}$$

For this example, i_L, v_{C_1}, and v_{C_2} have already been defined as state variables.

STEP 6: Substituting (10.28)–(10.30) into (10.16), (10.17), and (10.18) yields the desired state equation in matrix form:

$$\begin{bmatrix} \frac{dv_{C_1}}{dt} \\ \frac{dv_{C_2}}{dt} \\ \frac{di_L}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{1}{R_1 C_1} & 0 & -\frac{1}{C_1} \\ 0 & -\frac{1}{R_3 C_2} & \frac{1}{C_2} \\ \frac{1}{L} & -\frac{1}{L} & -\frac{R_2}{L} \end{bmatrix} \begin{bmatrix} v_{C_1} \\ v_{C_2} \\ i_L \end{bmatrix} + \begin{bmatrix} \frac{1}{R_1 C_1} & 0 \\ 0 & \frac{1}{C_2} \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_s \\ i_s \end{bmatrix} \tag{10.31}$$

which, as expected, is the same as (10.10) obtained previously by inspection.

As mentioned earlier, the selection of state variables is not unique. Instead of using capacitor voltages and inductor currents as state variables, basic quantities such as the capacitor charges and inductor fluxes may also be considered. If q_1, q_2, and ϕ_L are defined as state variables in Step 4, the inductive link current i_L and capacitive tree-branch voltages, v_{C_1} and v_{C_2}, can be solved from the inductor and capacitor equations in terms of state variables and possibly sources in Step 5 as

$$i_L = \frac{1}{L} \phi_L \tag{10.32}$$

$$v_{C_1} = \frac{1}{C_1} q_1 \tag{10.33}$$

$$v_{C_2} = \frac{1}{C_2} q_2 \tag{10.34}$$

Finally, state equations are obtained by substituting Eqs. (10.28)–(10.30) and (10.32)–(10.34) into (10.16)–(10.18) as

$$\begin{bmatrix} \frac{dq_1}{dt} \\ \frac{dq_2}{dt} \\ \frac{d\phi_L}{dt} \end{bmatrix} = \begin{bmatrix} -\frac{1}{R_1 C_1} & 0 & -\frac{1}{L} \\ 0 & -\frac{1}{R_3 C_2} & \frac{1}{L} \\ \frac{1}{C_1} & -\frac{1}{C_2} & -\frac{R_2}{L} \end{bmatrix} \begin{bmatrix} q_1 \\ q_2 \\ \Phi_L \end{bmatrix} + \begin{bmatrix} \frac{1}{R_1} & 0 \\ 0 & 1 \\ 0 & 0 \end{bmatrix} \begin{bmatrix} v_s \\ i_s \end{bmatrix} \tag{10.35}$$

In the systematic procedure outlined previously, it is assumed that the network exists with neither inductive cutsets nor capacitive loops so that the selection of proper tree is always guaranteed. For networks that do have these constraints, it is not possible to include all the capacitors in a tree without forming a closed path. Also, in order for a tree to contain all the nodes, some inductors will have to be included in a

tree. A tree that includes independent voltage sources, some resistors, and a maximum number of capacitors but no independent current sources is called a modified proper tree. In writing a state equation for such networks, the same systematic procedure can be applied with the selection of a modified proper tree. However, if capacitor tree-branch voltages and inductive link currents are defined as the state variables, the standard (**A, B, C, D**) description (10.14) and (10.15) may not exist. In fact, if inductive cutsets contain independent current sources and/or capacitive loops contain independent voltage sources, the derivative of these sources will appear in the state equation and the general equation is of the form

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} + \mathbf{B}_1\mathbf{u} + \mathbf{B}_2\dot{\mathbf{u}} \tag{10.36}$$

where $\mathbf{B}_1$ and $\mathbf{B}_2$ are $n \times r$ matrices and $\mathbf{A}$, $\mathbf{x}$, and $\mathbf{u}$ are defined as before. To recast (10.36) into the standard form, it is necessary to redefine.

$$\mathbf{z} = \mathbf{x} - \mathbf{B}_2\mathbf{u} \tag{10.37}$$

as new state variables. Substituting (10.37) into (10.36), yields

$$\dot{\mathbf{z}} = \mathbf{A}\mathbf{z} + \mathbf{B}\mathbf{u} \tag{10.38}$$

where

$$\mathbf{B} = \mathbf{B}_1 + \mathbf{A}\mathbf{B}_2 \tag{10.39}$$

It is noted from (10.37), the new state variables represent a linear combination of sources and capacitor voltages or inductor currents which, except for the mathematical convenience, may not have sound physical significance. To avoid such state variables and transformation (10.37), Step 4 of the systematic procedure described earlier needs to be modified. By defining state variables as the algebraic sum of capacitor charges in the fundamental cutset associated with each of the capacitor tree branches, and the algebraic sum of inductor fluxes in the fundamental loop associated with each of the inductive links, the resulting state equation will be in the standard form. The preceding generalizations are illustrated by the following two examples.

Example 5. As a simple illustration, consider the same circuits given in Figure 10.2, where the constant DC voltage source E is replaced by a time-varying source $e(t)$. It can easily be demonstrated that the equation describing the circuit now becomes

$$\frac{dv_{C_1}(t)}{dt} = -\frac{1}{R(C_1 + C_2)}v_{C_1} + \frac{1}{R(C_1 + C_2)}v_{in}(t) + \frac{C_2}{R(C_1 + C_2)}\frac{de(t)}{dt} \tag{10.40}$$

The preceding equation is the same as the state Eq. (10.4) with the exception of an additional term involving the first-order derivative of source $e(t)$. Equation (10.40) is clearly not the standard state equation described in (10.41) with capacitor voltage v_{C_1} defined as the state variable.

Example 6. As another illustration, consider the circuit shown in Figure 10.5 which consists of an inductive cutset $\{L_1, L_2, i_s\}$ and a capacitive loop (C_1, v_{s_2}, C_2). The state equations are determined from the systematic procedure by first using the transformation (10.37) and then by defining the algebraic sum of charges and fluxes as state variables.

STEP 1: The directed graph of the circuit is shown in Figure 10.6 where branches denoted by v_{s_1}, v_{s_2}, C_1, R_2, and L_2 are selected to form a modified proper tree.

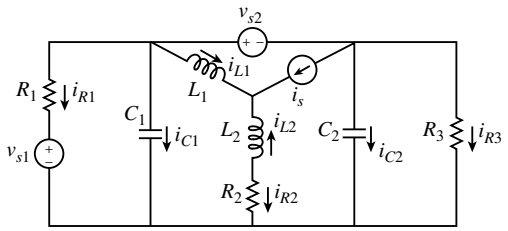


FIGURE 10.5 A circuit with a capacitive loop and an inductive cutset.

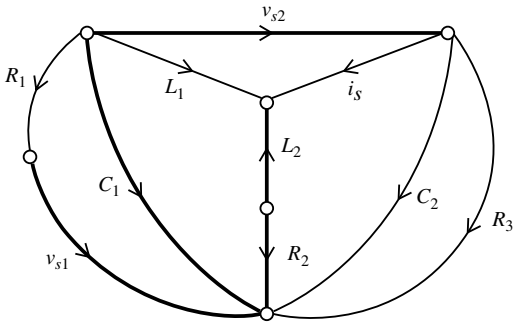


FIGURE 10.6 The directed graph associated with the circuit of Figure 10.5.

STEP 2: The fundamental cutset associated with C_1 consists of branches R_1, C_1, L_1, i_s, C_2, and R_3. Applying KCL to this cutset yields

$$i_{C_1} = -i_{R_1} - i_{L_1} - i_s - i_{C_2} - i_{R_3} \tag{10.41}$$

STEP 3: The fundamental loop equation associated with the inductive link L_1 is given by

$$v_{L_1} = v_{C_1} + v_{L_2} - v_{R_2} \tag{10.42}$$

where the link voltage v_{L_1} has been expressed in terms of tree-branch voltages.

STEP 4: In the first illustration, the tree-branch capacitor voltage v_{C_1} and the inductive link current i_{L_1} are defined as the state variables.

STEP 5: The branch relation and the remaining two fundamental equations are grouped into the following three sets:

Resistor equations:

$$v_{R_1} + v_{s_1} - v_{C_1} = 0 \tag{10.43}$$

$$v_{R_1} = R_1 i_{R_1} \tag{10.44}$$

$$i_{R_2} - i_{L_1} - i_s = 0 \tag{10.45}$$

$$v_{R_2} = R_2 i_{R_2} \tag{10.46}$$

$$v_{R_3} - v_{C_1} + v_{s_2} = 0 \tag{10.47}$$

$$v_{R_3} = R_3 i_{R_3} \tag{10.48}$$

Inductor equations:

$$i_{L_2} + i_{L_1} + i_s = 0 \tag{10.49}$$

$$\phi_{L_1} = L_1 i_{L_1} \quad \text{or} \quad v_{L_1} = L_1 \frac{di_{L_1}}{dt} \tag{10.50}$$

$$\phi_{L_2} = L_2 i_{L_2} \quad \text{or} \quad v_{L_2} = L_2 \frac{di_{L_2}}{dt} \tag{10.51}$$

Capacitor equations:

$$v_{C_2} - v_{C_1} + v_{s_2} = 0 \tag{10.52}$$

$$q_1 = C_1 v_{C_1} \quad \text{or} \quad i_{C_1} = C_1 \frac{dv_{C_1}}{dt} \tag{10.53}$$

$$q_2 = C_2 v_{C_2} \quad \text{or} \quad i_{C_2} = C_2 \frac{dv_{C_2}}{dt} \tag{10.54}$$

For this example, the nonstate variables are identified as i_{R1}, v_{R2}, i_{R3}, v_{L2}, and i_{C2}, from (10.41) and (10.42). These variables are now solved from the corresponding group of equations in terms of state variables and independent sources:

$$i_{R_1} = \frac{1}{R_1}\left(v_{C_1} - v_{s_1}\right) \tag{10.55}$$

$$v_{R_2} = R_2\left(i_{L_1} + i_s\right) \tag{10.56}$$

$$i_{R_3} = \frac{1}{R_3}\left(v_{C_1} - v_{s_2}\right) \tag{10.57}$$

$$v_{L_2} = -L_2 \frac{di_{L_1}}{dt} - L_2 \frac{di_s}{dt} \tag{10.58}$$

$$i_{C_2} = C_2 \frac{dv_{C_1}}{dt} - C_2 \frac{dv_{s_2}}{dt} \tag{10.59}$$

STEP 6: Assuming the existence of the first-order derivatives of sources with respect to time and substituting eqs. (10.50), (10.53), and (10.55)–(10.59) into (10.41) and (10.42) yields

$$\begin{bmatrix} \dfrac{dv_{C_1}}{dt} \\ \dfrac{di_{L_1}}{dt} \end{bmatrix} = \begin{bmatrix} -\dfrac{R_1+R_3}{R_1R_3(C_1+C_2)} & -\dfrac{1}{C_1+C_2} \\ \dfrac{1}{L_1+L_2} & -\dfrac{R_2}{L_1+L_2} \end{bmatrix} \begin{bmatrix} v_{C_1} \\ i_{L_1} \end{bmatrix}$$
$$+ \begin{bmatrix} \dfrac{1}{R_1(C_1+C_2)} & \dfrac{1}{R_3(C_1+C_2)} & -\dfrac{1}{(C_1+C_2)} \\ 0 & 0 & -\dfrac{R_2}{L_1+L_2} \end{bmatrix} \begin{bmatrix} v_{s_1} \\ v_{s_2} \\ i_s \end{bmatrix} \tag{10.60}$$
$$+ \begin{bmatrix} 0 & \dfrac{C_2}{(C_1+C_2)} & 0 \\ 0 & 0 & -\dfrac{L_2}{L_1+L_2} \end{bmatrix} \begin{bmatrix} \dfrac{dv_{s_1}}{dt} \\ \dfrac{dv_{s_2}}{dt} \\ \dfrac{di_s}{dt} \end{bmatrix}$$

Clearly, Eq. (10.60) is not in the standard form. Applying transformation (10.37) with $x_1 = v_{c1}$, $x_1 = i_{L2}$, $u_1 = v_{s1}$, $u_2 = v_{s2}$, and $u_3 = i_s$ gives the state equation in normal form

$$\begin{bmatrix} \dfrac{dz_1}{dt} \\ \dfrac{dz_2}{dt} \end{bmatrix} = \begin{bmatrix} -\dfrac{R_1+R_2}{R_1R_3(C_1+C_2)} & -\dfrac{1}{C_1+C_2} \\ \dfrac{1}{L_1+L_2} & -\dfrac{R_2}{L_1+L_2} \end{bmatrix} \begin{bmatrix} z_1 \\ z_2 \end{bmatrix}$$
$$+ \begin{bmatrix} \dfrac{1}{R_1(C_1+C_2)} & \dfrac{R_1C_1-R_3C_2}{R_1R_3(C_1+C_2)^2} & -\dfrac{L_1}{(L_1+L_2)(C_1+C_2)} \\ 0 & \dfrac{C_2}{(L_1+L_2)(C_1+C_2)} & -\dfrac{R_2L_1}{(L_1+L_2)^2} \end{bmatrix} \tag{10.61}$$
$$\times \begin{bmatrix} v_{s_1} \\ v_{s_2} \\ i_s \end{bmatrix}$$

where new state variables are defined as

$$\mathbf{z} = \begin{bmatrix} z_1 \\ z_2 \end{bmatrix} = \begin{bmatrix} v_{C_1} & -\dfrac{C_2}{C_1+C_2}v_{s_2} \\ i_{L_1} & +\dfrac{L_2}{L_1+L_2}i_s \end{bmatrix} \tag{10.62}$$

Alternatively, if the state variables are defined in Step 4 as

$$q_a = q_1 + q_2 \tag{10.63}$$

$$\phi_b = \phi_1 - \phi_2 \tag{10.64}$$

then Eqs. (10.41) and (10.42) become

$$\frac{dq_a}{dt} = \frac{dq_1}{dt} + \frac{dq_2}{dt} = -i_{R_1} - i_{L_1} - i_s - i_{R_3} \tag{10.65}$$

$$\frac{d\phi_b}{dt} = \frac{d\phi_1}{dt} - \frac{d\phi_2}{dt} = -v_{L_1} - v_{L_2} = v_{C_1} - v_{R_2} \tag{10.66}$$

respectively. In Step 5, the resistive link currents i_{R_1}, i_{R_3}, and the resistive tree-branch voltage V_{R_2} are solved from resistive eqs. (10.43)–(10.48) in terms of inductive link currents, capacitive tree-branch voltages, and independent sources. The results are those given in (10.55)–(10.57). By solving the inductor Eqs. (10.49), (10.50), and (10.64), inductive link current i_{L_1} is expressed as a function of state variables and independent sources:

$$i_{L_1} = \frac{1}{L_1 + L_2}\left(\phi_b - L_2 i_s\right) \tag{10.67}$$

Similarly, solving v_{C_1} from capacitor Eqs. (10.52)–(10.54), and (10.63), yields the capacitor tree-branch voltage

$$v_{C_1} = \frac{1}{C_1 + C_2}\left(q_a + C_2 v_{s_2}\right) \tag{10.68}$$

Finally, in Step 6, Eqs. (10.55)–(10.57), (10.67), and (10.68) are substituted into (10.65) and (10.66) to form the state equation in normal form:

$$\begin{bmatrix} \dfrac{dq_a}{dt} \\ \dfrac{d\phi_b}{dt} \end{bmatrix} = \begin{bmatrix} -\dfrac{R_1 + R_3}{R_1 R_3 (C_1 + C_2)} & -\dfrac{1}{L_1 + L_2} \\ \dfrac{1}{C_1 + C_2} & -\dfrac{R_2}{L_1 + L_2} \end{bmatrix} \begin{bmatrix} q_a \\ \phi_b \end{bmatrix} + \begin{bmatrix} \dfrac{1}{R_1} & \dfrac{R_1 C_1 - R_3 C_2}{R_1 R_3 (C_1 + C_2)} & -\dfrac{L_1}{(L_1 + L_2)} \\ 0 & \dfrac{C_2}{(C_1 + C_2)} & -\dfrac{R_2 L_1}{(L_1 + L_2)} \end{bmatrix} \begin{bmatrix} v_{s_1} \\ v_{s_2} \\ i_s \end{bmatrix} \tag{10.69}$$

10.3 Natural Response and State Transition Matrix

In the preceding section, the state-variable description has been presented for linear time-invariant circuits. The response of the circuit depends on the solution of the state equation. The behavior of the circuit due to any arbitrary input sources can easily be obtained once the zero-input response or the natural response of the circuit is known. In order to find its natural response, the homogeneous state equation of the circuit

$$\dot{\mathbf{x}} = \mathbf{A}\mathbf{x} \tag{10.70}$$

is considered, where independent source term $u(t)$ has been set equal to zero. The preceding state equation is analogous to the scalar equation

$$\dot{x} = ax \tag{10.71}$$

where the solution is given by

$$x(t) = e^{at}x(0) \tag{10.72}$$

for any arbitrary initial condition $x(0)$ given at $t = 0$, or

$$x(t) = e^{a(t-t_0)}x(t_0) \tag{10.73}$$

if the initial time is specified at $t = t_0$.

It is thus reasonable to assume a solution for (10.70) of the form

$$\mathbf{x}(t) = e^{(t-t_0)\lambda}\mathbf{p} \tag{10.74}$$

where λ is a scalar constant and $\mathbf{p}$ is a constant n-vector. Substituting (10.74) into (10.70) leads to

$$\mathbf{A}\mathbf{p} = \lambda\mathbf{p} \tag{10.75}$$

Therefore, (10.74) is a solution of (10.70) precisely when p is an eigenvector of $\mathbf{A}$ associated with the eigenvalue λ. For simplicity, it is assumed that $\mathbf{A}$ has n distinct eigenvalues $\lambda_1, \lambda_2, \ldots, \lambda_n$. Because the corresponding eigenvectors denoted by $\mathbf{p}_1, \mathbf{p}_2, \ldots, \mathbf{p}_n$ are linearly independent, the general solution of (10.70) can be uniquely written as a linear combination of n distinct normal modes of the form (10.74):

$$\mathbf{x}(t) = c_1 e^{(t-t_o)\lambda_1}\mathbf{p}_1 + c_2 e^{(t-t_0)\lambda_2}\mathbf{p}_2 + \cdots + c_n e^{(t-t_0)\lambda_n}\mathbf{p}_n \tag{10.76}$$

where $c_1, c_2, \ldots, c_n$ are n arbitrary constants determined by the given initial conditions. Specifically,

$$\mathbf{x}(t_0) = c_1\mathbf{p}_1 + c_2\mathbf{p}_2 + \cdots + c_n\mathbf{p}_n \tag{10.77}$$

The general solution (10.76) can also be written in the form

$$\mathbf{x}(t) = e^{(t-t_o)\mathbf{A}}\mathbf{x}(t_0) \tag{10.78}$$

where the exponential function of a matrix is defined by a power series:

$$\begin{aligned} e^{(t-t_o)\mathbf{A}} &= \mathbf{I} + (t-t_0)\mathbf{A} + \frac{(t-t_0)^2}{2!}\mathbf{A}^2 + \cdots \\ &= \sum_{k=0}^{\infty} \frac{(t-t_0)^k}{k!}\mathbf{A}^k \end{aligned} \tag{10.79}$$

In fact, taking the derivative of (10.78) with respect to t yields

$$\begin{aligned} \frac{d\mathbf{x}}{dt} &= \frac{d}{dt}\left[\mathbf{I} + (t-t_0)\mathbf{A} + \frac{(t-t_0)^2}{2!}\mathbf{A}^2 + \cdots\right]\mathbf{x}(t_0) \\ &= \left[\mathbf{A} + (t-t_0)\mathbf{A}^2 + \frac{(t-t_0)^2}{2!}\mathbf{A}^3 + \cdots\right]\mathbf{x}(t_0) \\ &= \mathbf{A}\left[\mathbf{I} + (t-t_0)\mathbf{A} + \frac{(t-t_0)^2}{2!}\mathbf{A}^2 + \cdots\right]\mathbf{x}(t_0) \\ &= \mathbf{A}e^{(t-t_0)\mathbf{A}}\mathbf{x}(t_0) = \mathbf{A}\mathbf{x}(t) \end{aligned} \tag{10.80}$$

Also, at $t = t_0$, (10.78) gives

$$\mathbf{x}(t_0) = \mathbf{I}\mathbf{x}(t_0) = \mathbf{x}(t_0) \tag{10.81}$$

Thus, expression (10.78) satisfies both eq. (10.70) and the initial conditions and hence is the unique solution. The matrix $e^{(t-t_0)\mathbf{A}}$, usually denoted by $\mathbf{\Phi}(t - t_0)$, is called the state transition matrix or the fundamental matrix of the circuit described by (10.70). The transition of the initial state $\mathbf{x}(t_0)$ to the state $\mathbf{x}(t)$ at any time t is thus governed by

$$\mathbf{x}(t) = \mathbf{\Phi}(t - t_0)\mathbf{x}(t_0) \tag{10.82}$$

where

$$\mathbf{\Phi}(t - t_0) = e^{(t-t_0)\mathbf{A}} \tag{10.83}$$

is an $n \times n$ matrix with the following properties:

$$\mathbf{\Phi}(t_0 - t_0) = \mathbf{\Phi}(0) = \mathbf{I} \tag{10.84}$$

$$\mathbf{\Phi}(t + \tau) = \mathbf{\Phi}(t)\mathbf{\Phi}(\tau) \tag{10.85}$$

$$\mathbf{\Phi}(t_2 - t_1)\mathbf{\Phi}(t_1 - t_0) = \mathbf{\Phi}(t_2 - t_0) \tag{10.86}$$

$$\mathbf{\Phi}(t_2 - t_1) = \mathbf{\Phi}^{-1}(t_1 - t_2) \tag{10.87}$$

$$\mathbf{\Phi}^{-1}(t) = \mathbf{\Phi}(-t) \tag{10.88}$$

Once the state transition matrix is known, the solution of the state equation can be obtained from (10.82). In general, it is rather difficult to obtain a closed-form solution from the infinite series representation of the state transition matrix. The formula given by (10.79) is useful only if numerical solution by digital computer is desired. Several methods are available for finding a closed form expression for $\mathbf{\Phi}(t - t_0)$. The relationship between solution (10.76) and the state transition matrix is first established.

For simplicity, let $t_0 = 0$. According to (10.82), the first column of $\mathbf{\Phi}(t)$ is the solution of the state equation generated by the initial condition

$$\mathbf{x}(0) = \mathbf{x}^{(1)}(0) = \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{10.89}$$

Indeed, if (10.89) is substituted into (10.82), then

$$\mathbf{x}(t) \triangleq \mathbf{x}^{(1)}(t) = \mathbf{\Phi}(t)\mathbf{x}^{(1)}(0) = \begin{bmatrix} \phi_{11} & \phi_{12} & \cdots & \phi_{1n} \\ \phi_{21} & \phi_{22} & \cdots & \phi_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ \phi_{n1} & \phi_{n2} & \cdots & \phi_{nn} \end{bmatrix} \begin{bmatrix} 1 \\ 0 \\ 0 \\ \vdots \\ 0 \end{bmatrix} = \begin{bmatrix} \phi_{11} \\ \phi_{21} \\ \vdots \\ \phi_{n1} \end{bmatrix} \tag{10.90}$$

which can be computed from (10.76) and the arbitrary constants $c_i \triangleq c_i^{(1)}$ for $i = 1, 2, \ldots, n$ are solved from (10.77). The first column of the state transition matrix is thus given by

$$\begin{bmatrix} \phi_{11} \\ \phi_{21} \\ \vdots \\ \phi_{n1} \end{bmatrix} = c_1^{(1)} e^{\lambda_1 t}\mathbf{p}_1 + c_2^{(1)} e^{\lambda_2 t}\mathbf{p}_2 + \cdots + c_n^{(1)} e^{\lambda_n t}\mathbf{p}_n \tag{10.91}$$

Instead of (10.89), if

$$\mathbf{x}(0) = \mathbf{x}^{(2)}(0) = \begin{bmatrix} 0 \\ 1 \\ 0 \\ \vdots \\ 0 \end{bmatrix} \tag{10.92}$$

is used, the arbitrary constants $c_1, c_2, \ldots, c_n$ denoted by $c_1^{(2)}, c_2^{(2)}, \ldots c_n^{(2)}$ are solved. Then, the second column of $\mathbf{\Phi}(t)$ is given

$$\begin{bmatrix} \phi_{12} \\ \phi_{22} \\ \vdots \\ \phi_{n2} \end{bmatrix} = c_1^{(2)} e^{\lambda_1 t}\mathbf{p}_1 + c_2^{(2)} e^{\lambda_2 t}\mathbf{p}_2 + \cdots + c_n^{(2)} e^{\lambda_n t}\mathbf{p}_n \tag{10.93}$$

In a similar manner, the remaining columns of $\mathbf{\Phi}(t)$ are determined.

The closed form expression for state transition matrix can also be obtained by means of a similarity transformation of the form

$$\mathbf{AP} = \mathbf{PJ}$$

or

$$\mathbf{J} = \mathbf{P}^{-1}\mathbf{AP} \tag{10.94}$$

where $\mathbf{P}$ is a nonsingular matrix. If the eigenvalues of $\mathbf{A}$, $\lambda_1, \lambda_2, \ldots, \lambda_n$, are assumed to be distinct, $\mathbf{J}$ is a diagonal matrix with eigenvalues on its main diagonal:

$$\mathbf{J} = \begin{bmatrix} \lambda_1 & 0 & \cdots & 0 \\ 0 & \lambda_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \lambda_n \end{bmatrix} \tag{10.95}$$

and

$$\mathbf{P} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 & \cdots & \mathbf{p}_n \end{bmatrix} \tag{10.96}$$

where $\mathbf{p}_i$'s, for i = 1,2,… *n*, are the corresponding eigenvectors associated with the eigenvalue λ_i, for i = 1,2,…,*n*. Substituting (10.94) into (10.83), the state transition matrix can now be written in the closed form

$$\begin{aligned}\mathbf{\Phi}(t-t_0) &= e^{(t-t_0)\mathbf{A}} = e^{(t-t_0)\mathbf{PJP}^{-1}} \\ &= \mathbf{P}e^{(t-t_0)J}\mathbf{P}^{-1}\end{aligned} \tag{10.97}$$

where

$$e^{(t-t_o)\mathbf{J}} = \begin{bmatrix} e^{(t-t_o)\lambda_1} & 0 & \cdots & 0 \\ 0 & e^{(t-t_o)\lambda_2} & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & e^{(t-t_o)\lambda_n} \end{bmatrix} \tag{10.98}$$

is a diagonal matrix.

In the more general case, where the **A** matrix has repeated eigenvalues, a diagonal matrix of the form (10.95) may not exist. However, it can be shown that any square matrix **A** can be transformed by a similarity transformation to the Jordan canonical form

$$\mathbf{J} = \begin{bmatrix} \mathbf{J}_1 & 0 & \cdots & 0 \\ 0 & \mathbf{J}_2 & \cdots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \cdots & \mathbf{J}_l \end{bmatrix} \tag{10.99}$$

where $\mathbf{J}_i$'s, for i = 1, 2,…, l are known as Jordan blocks. Assuming that **A** has m distinct eigenvalues, λ_i, with multiplicity r_i, for i = 1, 2, …, m, and $r_1 + r_2 + \cdots + r_m = n$. Associated with each λ_i there may exist several Jordan blocks. A Jordon block is a block diagonal matrix of order $k \times k (k \le r_i)$ with λ_i on its main diagonal, all 1's on the superdiagonal, and zeros elsewhere. In the special case when k = 1, the Jordan block reduces to a 1 × 1 scalar block with only one element λ_i.

In fact, the number of Jordan blocks associated with the eigenvalue λ_i is equal to the dimension of the null space of $(\lambda_i \mathbf{I} - \mathbf{A})$. For each $k \times k$ Jordan block $\mathbf{J}(k)$ associated with the eigenvalue λ_i of the form

$$\mathbf{J}(k) = \begin{bmatrix} \lambda_i & 1 & 0 & 0 & \cdots & 0 \\ 0 & \lambda_i & 1 & 0 & \cdots & 0 \\ \vdots & \vdots & \vdots & \vdots & \ddots & 1 \\ 0 & 0 & 0 & 0 & \cdots & \lambda_i \end{bmatrix} \tag{10.100}$$

the exponential function of $\mathbf{J}(k)$ takes the form

$$e^{(t-t_0)\mathbf{J}(\mathrm{k})} = \begin{bmatrix} 1 & t & \frac{t^2}{2!} & \cdots & \frac{t^{k-1}}{(k-1)!} \\ 0 & 1 & t & \cdots & \frac{t^{k-2}}{(k-2)!} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & 0 & \cdots & 1 \end{bmatrix} e^{(t-t_0)\lambda_i} \tag{10.101}$$

and the corresponding k columns of $\mathbf{P}$, known as the generalized eigenvectors, satisfy the equations

$$\begin{aligned}(\lambda_i\mathbf{I}-\mathbf{A})\mathbf{p}_i^{(1)} &= 0\\ (\lambda_i\mathbf{I}-\mathbf{A})\mathbf{p}_i^{(2)} &= -\mathbf{p}_i^{(1)}\\ &\vdots\\ (\lambda_i\mathbf{I}-\mathbf{A})\mathbf{p}_i^{(k)} &= -\mathbf{p}_i^{(k-1)}\end{aligned} \tag{10.102}$$

The closed form expression $\mathbf{\Phi}(t-t_0)$ for this general case now becomes

$$\mathbf{\Phi}(t-t_0) = \mathbf{P}e^{(t-t_0)J}\mathbf{P}^{-1} \tag{10.103}$$

where

$$e^{(t-t_0)\mathbf{J}} = \begin{bmatrix} e^{(t-t_0)\mathbf{J}_1} & 0 & \cdots & 0\\ 0 & e^{(t-t_0)\mathbf{J}_2} & \cdots & 0\\ \vdots & \vdots & \ddots & \vdots\\ 0 & 0 & \cdots & e^{(t-t_0)\mathbf{J}_l}\end{bmatrix} \tag{10.104}$$

and each of the $e^{(t-t_0)\mathbf{J}i}$, for $i = 1, 2, \ldots, l$, is of the form given in (10.101).

The third approach for obtaining closed form expression for the state transition matrix involves the Laplace transform technique. Taking the Laplace transform of (10.70) yields

$$s\mathbf{X}(s) - \mathbf{x}(0) = \mathbf{A}\mathbf{X}(s)$$

or

$$\mathbf{X}(s) = (s\mathbf{I} - \mathbf{A})^{-1}\,\mathbf{x}(0) \tag{10.105}$$

where $(s\mathbf{I} - \mathbf{A})^{-1}$ is known as the resolvent matrix. The time response

$$\mathbf{x}(t) = +^{-1}\left[(s\mathbf{I}-\mathbf{A})^{-1}\right]\mathbf{x}(0) \tag{10.106}$$

is obtained by taking the inverse Laplace transform of (10.105). It is observed by comparing (10.106) to (10.82) and (10.83) with $t_0 = 0$ that

$$\mathbf{\Phi}(t) = e^{t\mathbf{A}} = \mathcal{L}^{-1}\left[(s\mathbf{I}-\mathbf{A})^{-1}\right] \tag{10.107}$$

By way of illustration, the following example is considered. The state transition matrix is obtained by using each of the three approaches presented previously.

Example 7. Consider the parallel RLC circuit in Figure 10.7. The state equation of the circuit is obtained as

$$\begin{bmatrix}\dfrac{di_L}{dt}\\[2ex] \dfrac{dv_C}{dt}\end{bmatrix} = \begin{bmatrix} 0 & \dfrac{1}{L}\\[2ex] -\dfrac{1}{C} & -\dfrac{1}{RC}\end{bmatrix}\begin{bmatrix} i_L\\ v_C\end{bmatrix} + \begin{bmatrix} 0\\ \dfrac{1}{C}\end{bmatrix} i_s \tag{10.108}$$

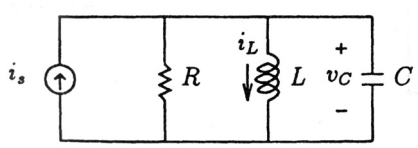


FIGURE 10.7 A parallel RLC circuit.

With $R = 2/3\ \Omega$, $L = 1\ H$, and $C = 1/2\ F$, the **A** matrix becomes

$$\mathrm{A} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \tag{10.109}$$

(a) *Normal Mode Approach:* The eigenvalues and the corresponding eigenvectors of the **A** are found to be

$$\lambda_1 = -1 \quad \lambda_2 = -2 \tag{10.110}$$

and

$$\mathbf{p}_1 = \begin{bmatrix} 1 \\ -1 \end{bmatrix}, \ \mathbf{p}_2 = \begin{bmatrix} 1 \\ -2 \end{bmatrix} \tag{10.111}$$

Therefore, the natural response of the circuit is given as a linear combination of the two distinct normal modes as

$$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = c_1 e^{-t} \begin{bmatrix} 1 \\ -1 \end{bmatrix} + c_2 e^{-2t} \begin{bmatrix} 1 \\ -2 \end{bmatrix} \tag{10.112}$$

When evaluated at $t = 0$, (10.112) becomes

$$\begin{bmatrix} i_L(0) \\ v_C(0) \end{bmatrix} = \begin{bmatrix} c_1 + c_2 \\ -c_1 - 2c_2 \end{bmatrix} \tag{10.113}$$

In order to find the first column of $\mathbf{\Phi}(t)$, it is assumed that

$$\begin{bmatrix} i_L(0) \\ v_C(0) \end{bmatrix} = \begin{bmatrix} 1 \\ 0 \end{bmatrix} \tag{10.114}$$

With this initial condition, the solution of (10.113) becomes

$$c_1 \triangleq c_1^{(1)} = 2 \quad \text{and} \quad c_2 \triangleq c_2^{(1)} = -1 \tag{10.115}$$

Substituting (10.115) into (10.112) results in the first column of $\mathbf{\Phi}(t)$:

$$\begin{bmatrix} \phi_{11} \\ \phi_{21} \end{bmatrix} = \begin{bmatrix} 2e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} \end{bmatrix} \tag{10.116}$$

Similarly, for

$$\begin{bmatrix} i_L(0) \\ v_C(0) \end{bmatrix} = \begin{bmatrix} 0 \\ 1 \end{bmatrix} \tag{10.117}$$

constants c_1 and c_2 are solved from (10.113) to give

$$c_1 \triangleq c_2^{(2)} = 1 \text{ and } c_2 \triangleq c_2^{(2)} = -1 \tag{10.118}$$

The second column of $\boldsymbol{\Phi}(t)$:

$$\begin{bmatrix} \phi_{12} \\ \phi_{22} \end{bmatrix} = \begin{bmatrix} e^{-t} - e^{-2t} \\ -e^{-t} + 2e^{-2t} \end{bmatrix} \tag{10.119}$$

is obtained by substituting (10.118) into (10.112). Combining (10.116) and (10.119) yields the state transition matrix in closed form

$$\boldsymbol{\Phi}(t) = \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix} \tag{10.120}$$

(**b**) *Similarity Transformation Method:* The eigenvalues are distinct, so the nonsingular transformation **P** is constructed from (10.96) by the eigenvectors of **A**:

$$\mathbf{P} = \begin{bmatrix} \mathbf{p}_1 & \mathbf{p}_2 \end{bmatrix} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \tag{10.121}$$

with

$$\mathbf{P}^{-1} = \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \tag{10.122}$$

Substituting λ_1, λ_2, and **P** into (10.97) and (10.98) yields the desired state transition matrix

$$\boldsymbol{\Phi}(t) = \mathbf{P}e^{\mathbf{J}}\mathbf{P}^{-1} = \begin{bmatrix} 1 & 1 \\ -1 & -2 \end{bmatrix} \begin{bmatrix} e^{-t} & 0 \\ 0 & e^{-2t} \end{bmatrix} \begin{bmatrix} 2 & 1 \\ -1 & -1 \end{bmatrix} \tag{10.123}$$

$$= \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix}$$

which is in agreement with (10.120).

(**c**) *Laplace Transform Technique:* The state transition matrix can also be computed in the frequency domain from (10.107). The resolvent matrix is

$$\begin{aligned} (s\mathbf{I} - \mathbf{A})^{-1} &= \begin{bmatrix} s & -1 \\ 2 & s+3 \end{bmatrix}^{-1} \\ &= \begin{bmatrix} \dfrac{s+3}{(s+1)(s+2)} & \dfrac{1}{(s+1)(s+2)} \\ \dfrac{-2}{(s+1)(s+2)} & \dfrac{s}{(s+1)(s+2)} \end{bmatrix} \end{aligned} \tag{10.124}$$

$$= \begin{bmatrix} \dfrac{2}{s+1} - \dfrac{1}{s+2} & \dfrac{1}{s+1} - \dfrac{1}{s+2} \\ \dfrac{2}{s+1} + \dfrac{2}{s+2} & -\dfrac{1}{s+1} + \dfrac{2}{s+2} \end{bmatrix}$$

where partial-fraction expansion has been applied. Taking the inverse Laplace transform of (10.124) yields the same closed form expression as given previously in (10.120) for $\mathbf{\Phi}(t)$.

10.4 Complete Response

When independent sources are present in the circuit, the complete response depends on the initial states of the circuits as well as the input sources. It is well known that the complete response is the sum of the zero-input (or natural) response and the zero-state (or forced) response and satisfies the nonhomogeneous state equation

$$\dot{\mathbf{x}}(t) = \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \tag{10.125}$$

subject to the given initial condition $\mathbf{x}(t_0) = \mathbf{x}_0$. Equation (10.125) is again analogous to the scalar equation

$$\dot{x}(t) = ax(t) + bu(t) \tag{10.126}$$

which has the unique solution of the form

$$x(t) = e^{(t-t_0)a}x(t_0) + \int_{t_0}^{t} e^{(t-\tau)a}bu(\tau)\,d\tau \tag{10.127}$$

It is thus assumed that the solution to the state equation is given by

$$\begin{aligned} \mathbf{x}(t) &= e^{(t-t_0)\mathbf{A}}\mathbf{x}(t_0) + \int_{t_0}^{t} e^{(t-\tau)\mathbf{A}}\mathbf{B}\mathbf{u}(\tau)d\tau \\ &= \mathbf{\Phi}(t-t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t-\tau)\mathbf{B}\mathbf{u}(\tau)d\tau \end{aligned} \tag{10.128}$$

Indeed, one can show by direct substitution that (10.128) satisfies the state Eq. (10.125). Differentiating both sides of (10.128) with respect to t yields

$$\begin{aligned} \dot{\mathbf{x}}(t) &= \frac{d}{dt}\mathbf{\Phi}(t-t_0)\mathbf{x}(t_0) + \frac{d}{dt}\int_{t_0}^{t} \mathbf{\Phi}(t-\tau)\mathbf{B}\mathbf{u}(\tau)d\tau \\ &= \mathbf{A}\mathbf{\Phi}(t-t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \frac{d}{dt}\mathbf{\Phi}(t-\tau)\mathbf{B}\mathbf{u}(\tau)d\tau + \mathbf{\Phi}(t-t)\mathbf{B}\mathbf{u}(t) \\ &= \mathbf{A}\mathbf{\Phi}(t-t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{A}\mathbf{\Phi}(t-\tau)\mathbf{B}\mathbf{u}(\tau)d\tau + \mathbf{B}\mathbf{u}(t) \\ &= \mathbf{A}\left[\mathbf{\Phi}(t-t_0)\mathbf{x}(t_0) + \int_{t_0}^{t} \mathbf{\Phi}(t-\tau)\mathbf{B}\mathbf{u}(\tau)d\tau\right] + \mathbf{B}\mathbf{u}(t) \\ &= \mathbf{A}\mathbf{x}(t) + \mathbf{B}\mathbf{u}(t) \end{aligned} \tag{10.129}$$

Also, at $t = t_0$, (10.128) becomes

$$\begin{aligned} \mathbf{x}(t_0) &= \mathbf{\Phi}(t_0 - t_0)\mathbf{x}(t_0) + \int_{t_0}^{t_0} \mathbf{\Phi}(t_0 - \tau)\mathbf{B}\mathbf{u}(\tau)d\tau \\ &= \mathbf{I}\mathbf{x}(t_0) + \mathbf{0} = \mathbf{x}(t_0) \end{aligned} \tag{10.130}$$

The assumed solution (10.128) thus satisfies both the state Eq. (10.125) and the given initial condition. Hence, $\mathbf{x}(t)$ as given by (10.128) is the unique solution.

It is observed from (10.128) that if $\mathbf{u}(t)$ is set to zero, the solution reduces to the zero-input response or the natural response given in (10.82). On the other hand, if the original circuit is relaxed, i.e., $\mathbf{x}(t_0) = 0$, the solution represented by the convolution integral, the second term on the right-hand side of (10.128), is the forced response on the zero-state response. Thus, Eq. (10.128) verifies the fact that the complete response is the sum of the zero-input response and the zero-state response. The previous result is illustrated by means of the following example.

Example 8. Consider again the same circuit given in Example 7, where the input current source is assumed to be a unit step function applied to the circuit at $t = 0$.

The state equation of the circuit is found from (10.108) to be

$$\begin{bmatrix} \dfrac{di_L}{dt} \\ \dfrac{dv_C}{dt} \end{bmatrix} = \begin{bmatrix} 0 & 1 \\ -2 & -3 \end{bmatrix} \begin{bmatrix} i_L \\ v_C \end{bmatrix} + \begin{bmatrix} 0 \\ 2 \end{bmatrix} i_s(t) \tag{10.131}$$

where the state transition matrix $\mathbf{\Phi}(t)$ is given in (10.120).

The zero-state response for $t > 0$ is obtained by evaluating the convolution integral indicated in (10.128):

$$\begin{aligned} \int_0^t \mathbf{\Phi}(t-\tau)\mathbf{B}\mathbf{u}(\tau)d\tau &= \int_0^t \begin{bmatrix} 2e^{-(t-\tau)} - e^{-2(t-\tau)} & e^{-(t-\tau)} - e^{-2(t-\tau)} \\ -2e^{-(t-\tau)} + 2e^{-2(t-\tau)} & -e^{-(t-\tau)} + 2e^{-2(t-\tau)} \end{bmatrix} \begin{bmatrix} 0 \\ 2 \end{bmatrix} d\tau \\ &= 2\int_0^t \begin{bmatrix} e^{-(t-\tau)} - e^{-2(t-\tau)} \\ -e^{-(t-\tau)} + 2e^{-2(t-\tau)} \end{bmatrix} d\tau \\ &= \begin{bmatrix} 1 - 2e^{-t} + e^{-2t} \\ 2e^{-t} - 2e^{-2t} \end{bmatrix} \end{aligned} \tag{10.132}$$

By adding the zero-input response represented by $\mathbf{\Phi}(t)\mathbf{x}(0)$ to (10.132), the complete response for any given initial condition $\mathbf{x}(0)$ becomes

$$\begin{bmatrix} i_L(t) \\ v_C(t) \end{bmatrix} = \begin{bmatrix} 2e^{-t} - e^{-2t} & e^{-t} - e^{-2t} \\ -2e^{-t} + 2e^{-2t} & -e^{-t} + 2e^{-2t} \end{bmatrix} \begin{bmatrix} i_L(0) \\ v_C(0) \end{bmatrix} + \begin{bmatrix} 1 - 2e^{-t} + 2e^{-2t} \\ 2e^{-t} - 2e^{-2t} \end{bmatrix} \tag{10.133}$$

for $t > 0$.

References

[1] T. C. Chen, *Linear System Theory and Design*, New York: Holt, Rinehart & Winston, 1970.

[2] W. K. Chen, *Linear Networks and Systems*, Monterey, CA: Brooks/Cole Engineering Division, 1983.

[3] L. O. Chua and P. M. Lin, *Computer-Aided Analysis of Electronics Circuits: Algorithms and Computational Techniques*, Englewood Cliffs, NJ: Prentice Hall, 1969.

[4] P. M. DeRusso, R. J. Roy, and C. M. Close, *State Variables for Engineers*, New York: John Wiley & Sons, 1965.

[5] C. A. Desoer and E. S. Kuh, *Basic Circuit Theory*, New York: McGraw-Hill, 1969.

[6] B. C. Kuo, *Linear Networks and Systems*, New York: McGraw-Hill, 1967.

[7] K. Ogata, *State Space Analysis of Control Systems*, Englewood Cliffs, NJ: Prentice Hall, 1967.

[8] R. A. Rohrer, Circuit *Theroy*: *An Introduction to the State Variable Approach*, New York: McGraw-Hill, 1970.
[9] D. G. Schultz and J. L. Melsa, *State Functions and Linear Control Systems*, New York: McGraw-Hill, 1967.
[10] T. E. Stern, *Theory of Nonlinear Networks and Systems: An Introduction*, Reading MA: Addison-Wesley, 1965.
[11] L. K. Timothy and B. E. Bona, *State Space Analysis: An Introduction*, New York: McGraw Hill, 1968.
[12] L. A. Zadeh and C. A. Desoer, *Linear System Theory*, New York: McGraw-Hill, 1963.

11

Feedback Amplifier Theory

John Choma, Jr.
University of Southern California

11.1 Introduction

Feedback, whether intentional or parasitic, is pervasive of all electronic circuits and systems. In general, feedback is comprised of a subcircuit that allows a fraction of the output signal of an overall network to modify the effective input signal in such a way as to produce a circuit response that can differ substantially from the response produced in the absence of such feedback. If the magnitude and relative phase angle of the fed back signal decreases the magnitude of the signal applied to the input port of an amplifier, the feedback is said to be *negative* or **degenerative**. On the other hand, *positive* (or **regenerative**) feedback, which gives rise to oscillatory circuit responses, is the upshot of a feedback signal that increases the magnitude of the effective input signal. Because negative feedback produces stable circuit responses, the majority of all intentional feedback architectures is degenerative [1], [2]. However, parasitic feedback incurred by the energy storage elements associated with circuit layout, circuit packaging, and second-order high-frequency device phenomena often degrades an otherwise degenerative feedback circuit into either a potentially regenerative or severely underdamped network.

Intentional degenerative feedback applied around an analog network produces four circuit performance benefits. First, negative feedback desensitizes the gain of an **open-loop amplifier** (an amplifier implemented without feedback) with respect to variations in circuit element and active device model parameters. This desensitization property is crucial in view of parametric uncertainties caused by aging phenomena, temperature variations, biasing perturbations, and nonzero fabrication and manufacturing tolerances. Second, and principally because of the foregoing desensitization property, degenerative feedback reduces the dependence of circuit responses on the parameters of inherently nonlinear active devices, thereby reducing the total harmonic distortion evidenced in open loops. Third, negative feedback broadbands the dominant pole of an open-loop amplifier, thereby affording at least the possibility of a closed-loop network with improved high-frequency performance. Finally, by modifying the driving-point input and output impedances of the open-loop circuit, negative feedback provides a convenient vehicle for implementing voltage buffers, current buffers, and matched interstage impedances.

The disadvantages of negative feedback include gain attenuation, a closed-loop configuration that is disposed to potential instability, and, in the absence of suitable frequency compensation, a reduction in the open-loop gain-bandwidth product. In uncompensated feedback networks, open-loop amplifier gains are reduced in almost direct proportion to the amount by which closed-loop amplifier gains are desensitized with respect to open-loop gains. Although the 3-dB bandwidth of the open-loop circuit is increased by a factor comparable to that by which the open-loop gain is decreased, the closed-loop gain-bandwidth product resulting from uncompensated degenerative feedback is never greater than that of the open-loop configuration [3]. Finally, if feedback is incorporated around an open-loop amplifier that does not have a dominant pole [4], complex conjugate closed-loop poles yielding nonmonotonic frequency responses are likely. Even positive feedback is possible if substantive negative feedback is applied around an open-loop amplifier for which more than two poles significantly influence its frequency response.

Although the foregoing detail is common knowledge deriving from Bode's pathfinding disclosures [5], most circuit designers remain uncomfortable with analytical procedures for estimating the frequency responses, I/O impedances, and other performance indices of practical feedback circuits. The purposes of this section are to formulate systematic feedback circuit analysis procedures and ultimately, to demonstrate their applicability to six specific types of commonly used feedback architectures. Four of these feedback types, the series-shunt, shunt-series, shunt-shunt, and series-series configurations, are single-loop architectures, while the remaining two types are the series-series/shunt-shunt and series-shunt/shunt-series dual-loop configurations.

11.2 Methods of Analysis

Several standard techniques are used for analyzing linear feedback circuits [6]. The most straightforward of these entails writing the Kirchhoff equilibrium equations for the small-signal model of the entire feedback system. This analytical tack presumably leads to the idealized feedback circuit block diagram abstracted in Figure 11.1. In this model, the circuit voltage or current response, X_R, is related to the source current or voltage excitation, X_S, by

$$G_{cl} \triangleq \frac{X_R}{X_S} = \frac{G_o}{1 + f G_o} \equiv \frac{G_o}{1 + T} \tag{11.1}$$

where G_{cl} is the closed-loop gain of the feedback circuit, the feedback factor f is the proportion of circuit response fed back for antiphase superposition with the source signal, and G_o represents the open-loop gain. The product fG_o is termed the loop gain T.

Equation (11.1) demonstrates that, for loop gains with magnitudes that are much larger than one, the closed-loop gain collapses to $1/f$, which is independent of the open-loop gain. To the extent that the

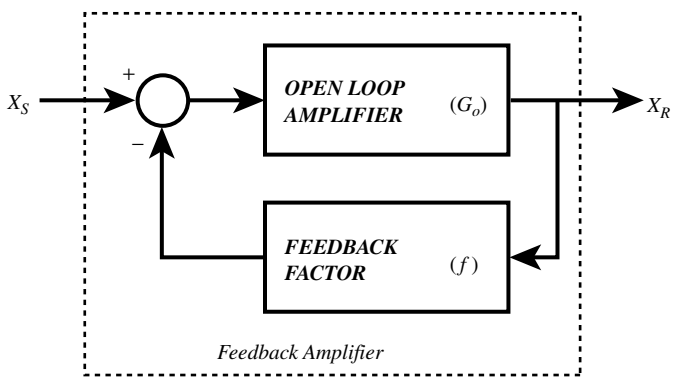


FIGURE 11.1 Block diagram model of a feedback network.

open-loop amplifier, and not the feedback subcircuit, contains circuit elements and other parameters that are susceptible to modeling uncertainties, variations in the fabrication of active and passive elements, and nonzero manufacturing tolerances, large loop gain achieves a desirable parametric desensitization. Unfortunately, the determination of G_o and f directly from the Kirchhoff relationships is a nontrivial task, especially because G_o is rarely independent of f in practical electronics. Moreover, (11.1) does not illuminate the manner in which the loop gain modifies the driving-point input and output impedances of the open-loop amplifier.

A second approach to feedback network analysis involves modeling the open-loop, feedback, and overall closed-loop networks by a homogeneous set of two-port parameters [7]. When the two-port parameter model is selected judiciously, the two-port parameters for the closed-loop network derive from a superposition of the respective two-port parameters of the open-loop and feedback subcircuits. Given the resultant parameters of the closed-loop circuit, standard formulas can then be exploited to evaluate closed-loop values of the circuit gain and the driving-point input and output impedances.

Unfortunately, several limitations plague the utility of feedback network analysis predicated on two-port parameters. First, the computation of closed-loop two-port parameters is tedious if the open-loop configuration is a multistage amplifier, or if multiloop feedback is utilized. Second, the two-loop method of feedback circuit analysis is straightforwardly applicable to only those circuits that implement **global feedback** (feedback applied from output port to input port). Many single-ended feedback amplifiers exploit only **local feedback**, wherein a fraction of the signal developed at the output port is fed back to a terminal pair other than that associated with the input port. Finally, the appropriate two-port parameters of the open-loop amplifier can be superimposed with the corresponding parameter set of the feedback subcircuit if and only if the Brune condition is satisfied [8]. This requirement mandates equality between the preconnection and postconnection values of the two-port parameters of open-loop and feedback cells, respectively. The subject condition is often not satisfied when the open-loop amplifier is not a simple three-terminal two-port configuration.

The third method of feedback circuit analysis exploits Mason's signal flow theory [9–11]. The circuit level application of this theory suffers few of the shortcomings indigenous to block diagram and two-port methods of feedback circuit analysis [12]. Signal flow analyses applied to feedback networks efficiently express I/O transfer functions, driving-point input impedances, and driving-point output impedances in terms of an arbitrarily selected critical or reference circuit parameters, say P.

An implicit drawback of signal flow methods is the fact that unless P is selected to be the feedback factor f, which is not always transparent in feedback architectures, expressions for the loop gain and the open-loop gain of feedback amplifiers are obscure. However, by applying signal flow theory to a feedback circuit model engineered from insights that derive from the results of two-port network analyses, the feedback factor can be isolated. The payoff of this hybrid analytical approach includes a conventional block diagram model of the I/O transfer function, as well as convenient mathematical models for evaluating the closed-loop driving-point input and output impedances. Yet, another attribute of hybrid methods of feedback circuit analysis is its ability to delineate the cause, nature, and magnitude of the feedforward transmittance produced by interconnecting a certain feedback subcircuit to a given open-loop amplifier. This information is crucial in feedback network design because feedforward invariably decreases gain and often causes undesirable phase shifts that can lead to significantly underdamped or unstable closed-loop responses.

11.3 Signal Flow Analysis

Guidelines for feedback circuit analysis by hybrid signal flow methods can be established with the aid of Figure 11.2 [13]. Figure 11.2(a) depicts a linear network whose output port is terminated in a resistance, R_L. The output signal variable is the voltage V_O, which is generated in response to an input port signal whose Thévenin voltage and resistance are respectively, V_S and R_S. Implicit to the linear network is a current-controlled voltage source (CCVS) Pi_b, with a value that is directly proportional to the indicated network branch current i_b. The problem at hand is the deduction of the voltage gain $G_v(R_S, R_L) = V_O/V_S$,

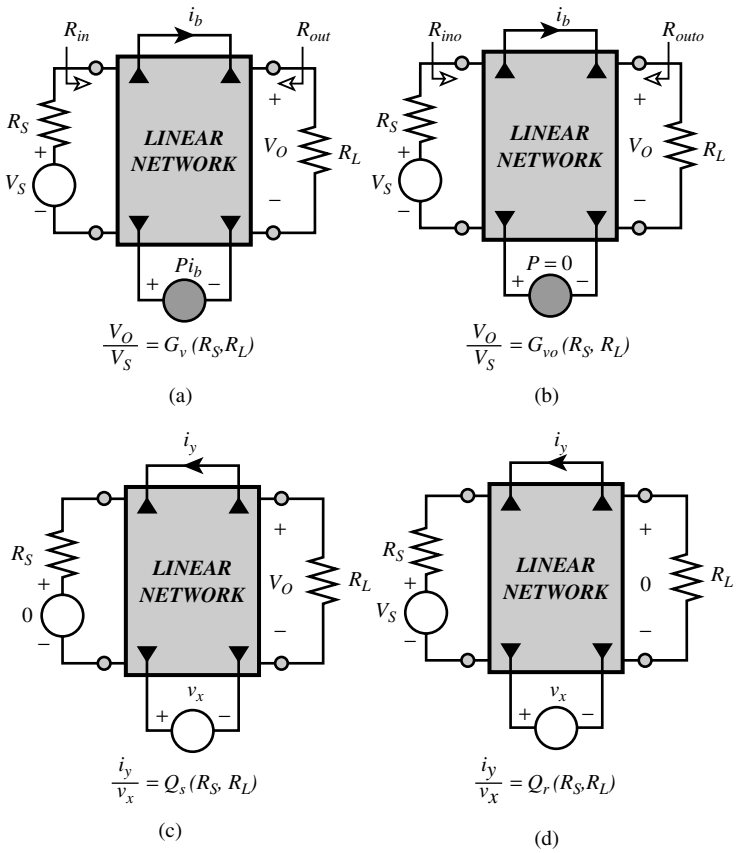


FIGURE 11.2 (a) Linear network with an identified critical parameter P. (b) Model for calculating the $P = 0$ value of voltage gain. (c) The return ratio with respect to P is $PQ_s(R_S, R_L)$. (d) The null return ratio with respect to P is $PQ_r(R_S, R_L)$.

the driving-point input resistance (or impedance) R_{in}, and the driving-point output resistance (or impedance) R_{out}, as explicit functions of the critical transimpedance parameter P. Although the following systematic procedure is developed in conjunction with the diagram in Figure 11.2, with obvious changes in notation, it is applicable to determining any type of transfer relationship for any linear network in terms of any type of reference parameter [14].

1. Set $P = 0$, as depicted in Figure 11.2(b), and compute the resultant voltage gain $G_{vo}(R_S, R_L)$, where the indicated notation suggests an anticipated dependence of gain on source and load resistances. Also, compute the corresponding driving-point input and output resistances R_{in}, and R_{out}, respectively. In this case, the "critical" parameter P is associated with a controlled voltage source. Accordingly, $P = 0$ requires that the branch containing the controlled source be supplanted by a short circuit. If, for example, P is associated with a controlled current source, $P = 0$ mandates the replacement of the controlled source by an open circuit.
2. Set the Thévenin source voltage V_S to zero, and replace the original controlled voltage source Pi_b by an independent voltage source of symbolic value, v_x. Then, calculate the ratio, i_y/v_x, where, as illustrated in Figure 11.2(c), i_y flows in the branch that originally conducts the controlling current i_b. Note, however, that the reference polarity of i_y is opposite to that of i_b. The computed transfer function i_y/v_x is denoted by $Q_S(R_S, R_L)$. This transfer relationship, which is a function of the source and load resistances, is used to determine the *return ratio* $T_s(P, R_S, R_L)$ with respect to parameter P of the original network. In particular,

$$T_s(P, R_S, R_L) = PQ_s(R_S, R_L) \tag{11.2}$$

If P is associated with a controlled current source, the controlled generator Pi_b is replaced by a current source of value i_x. If the controlling variable is a voltage, instead of a current, the ratio v_y/v_x, is computed, where v_y, where the polarity is opposite to that of the original controlling voltage, is the voltage developed across the controlling branch.

3. The preceding computational step is repeated, but instead of setting V_S to zero, the output variable, which is the voltage V_O in the present case, is nulled, as indicated in Figure 11.2(d). Let the computed ratio i_y/v_x, be symbolized as $Q_r(R_S, R_L)$. In turn, the *null return ratio* $T_r(P, R_S, R_L)$, with respect to parameter P is

$$T_r(P, R_S, R_L) = PQ_r(R_S, R_L) \tag{11.3}$$

4. The desired voltage gain $G_v(R_S, R_L)$, of the linear network undergoing study can be shown to be [5, 12]

$$G_v(R_S, R_L) = \frac{V_O}{V_S} = G_{vo}(R_S, R_L)\left[\frac{1 + PQ_r(R_S, R_L)}{1 + PQ_s(R_S, R_L)}\right] \tag{11.4}$$

5. Given the function $Q_s(R_S, R_L)$, the driving-point input and output resistances follow straightforwardly from [12]

$$R_{\text{in}} = R_{\text{ino}}\left[\frac{1 + PQ_s(0, R_L)}{1 + PQ_s(\infty, R_L)}\right] \tag{11.5}$$

$$R_{\text{out}} = R_{\text{outo}}\left[\frac{1 + PQ_s(R_S, 0)}{1 + PQ_s(R_S, \infty)}\right] \tag{11.6}$$

An important special case entails a controlling electrical variable i_b associated with the selected parameter P that is coincidentally the voltage or current output of the circuit under investigation. In this situation, a factor P of the circuit response is fed back to the port (not necessarily the input port) defined by the terminal pair across which the controlled source is incident. When the controlling variable i_b is the output voltage or current of the subject circuit $Q_r(R_S, R_L)$, which is evaluated under the condition of a nulled network response, is necessarily zero. With $Q_r(R_S, R_L) = 0$, the algebraic form of (11.4) is identical to that of (11.1), where the loop gain T is the return ratio with respect to parameter P; that is,

$$PQ_s(R_S, R_L)\Big|_{Q_r(R_S, R_L)=0} = T \tag{11.7}$$

Moreover, a comparison of (11.4) to (11.l) suggests that $G_v(R_S, R_L)$ symbolizes the closed-loop gain of the circuit, $G_{vo}(R_S, R_L)$ represents the corresponding open-loop gain, and the circuit feedback factor f is

$$f = \frac{PQ_S(R_S, R_L)}{G_{vo}(R_S, R_L)} \tag{11.8}$$

11.4 Global Single-Loop Feedback

Consider the global feedback scenario illustrated in Figure 11.3(a), in which a fraction P of the output voltage V_O is fed back to the voltage-driven input port. Figure 11.3(b) depicts the model used to calculate

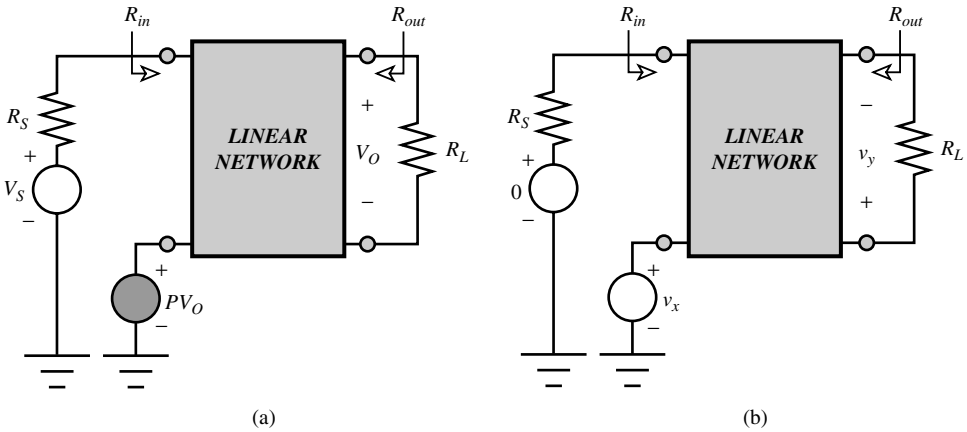


FIGURE 11.3 (a) Voltage-driven linear network with global voltage feedback. (b) Model for the calculation of loop gain.

the return ratio $Q_s(R_S, R_L)$, where, in terms of the branch variables in the schematic diagram, $Q_s(R_S, R_L) = v_y/v_x$. An inspection of this diagram confirms that the transfer function v_y/v_x, is identical to the $P = 0$ value of the gain V_O/V_S, which derives from an analysis of the structure in Figure 11.3(a). Thus, for global voltage feedback in which a fraction of the output voltage is fed back to a voltage-driven input port, $Q_s(R_S, R_L)$ is the open-loop voltage gain; that is, $Q_s(R_S, R_L) \equiv G_{vo}(R_S, R_L)$. It follows from (11.8) that the feedback factor f is identical to the selected critical parameter P. Similarly, for the global current feedback architecture of Figure 11.4(a), in which a fraction P of the output current, I_O, is feed back to the current-driven input port $f = P$. As implied by the model of Figure 11.4(b), $Q_s(R_S, R_L) \equiv G_{io}(R_S, R_L)$, the open-loop current gain.

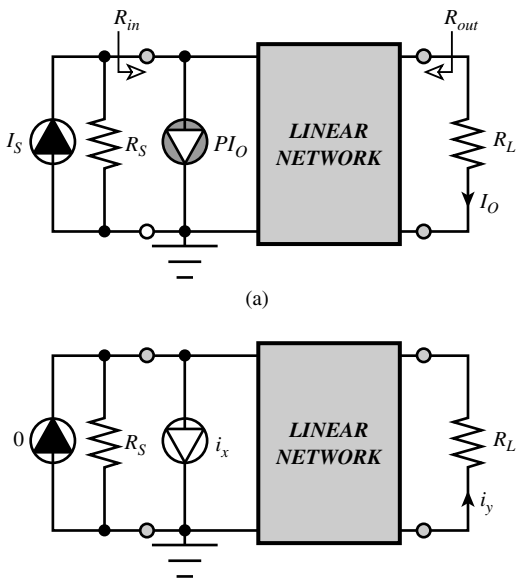


FIGURE 11.4 (a) Current-driven linear network with global current feedback. (b) Model for the calculation of loop gain.

Driving-Point I/O Resistances

Each of the two foregoing circuit architectures has a closed-loop gain where the algebraic form mirrors (11.1). It follows that for sufficiently large loop gain [equal to either $PG_{vo}(R_S, R_L)$ or $PG_{io}(R_S, R_L)$], the closed-loop gain approaches $(1/P)$ and is therefore desensitized with respect to open-loop gain parameters. However, such a desensitization with respect to the driving-point input and output resistances (or impedances) cannot be achieved. For the voltage feedback circuit in Figure 11.3(a), $Q_s(\infty, R_L)$, is the $R_S = \infty$ value, $G_{vo}(R_S,R_L)$, of the open-loop voltage gain. This particular open-loop gain is zero, because $R_S = \infty$ decouples the source voltage from the input port of the amplifier. On the other hand, $Q_s(0, R_L)$ is the $R_S = 0$ value, $G_{vo}(0, R_L)$, of the open-loop voltage gain. This gain is at least as large as $G_{vo}(R_S, R_L)$, since a short circuited Thévenin source resistance implies lossless coupling of the Thévenin signal to the amplifier input port. Recalling (11.5), the resultant driving-point input resistance of the voltage feedback amplifier is

$$R_{\text{in}} = R_{\text{in}o}\left[1 + PG_{vo}\left(0, R_L\right)\right] \geq R_{\text{in}o}\left[1 + PG_{vo}\left(R_S, R_L\right)\right] \tag{11.9}$$

which shows that the closed-loop driving-point input resistance is larger than its open-loop counterpart and is dependent on open-loop voltage gain parameters.

Conversely, the corresponding driving-point output resistance in Figure 11.3(a) is smaller than the open-loop output resistance and approximately inversely proportional to the open-loop voltage gain. These assertions derive from the facts that $Q_s(R_S, 0)$ is the $R_L = 0$ value of the open-loop voltage gain $G_{vo}(R_S, R_L)$. Because $R_L = 0$ corresponds to the short-circuited load resistance, $G_{vo}(R_S, 0) = 0$. In contrast, $Q_s(R_S, \infty)$, is the $R_L = \infty$ value, $G_{vo}(R_S, \infty)$, of the open-loop gain, which is a least as large as $G_{vo}(R_S, R_L)$. By (11.6),

$$R_{\text{out}} = \frac{R_{\text{out}o}}{1 + PG_{vo}\left(R_S, \infty\right)} \leq \frac{R_{\text{out}o}}{1 + PG_{vo}\left(R_S, R_L\right)} \tag{11.10}$$

Similarly, the driving-point input and output resistances of the global current feedback configuration of Figure 11.4(a) are sensitive to open-loop gain parameters. In contrast to the voltage amplifier of Figure 11.3(a), the closed-loop, driving-point input resistance of current amplifier is smaller than its open-loop value, while the driving-point output resistance is larger than its open-loop counterpart. Noting that the open-loop current gain $G_{io}(R_S, R_L)$ is zero for both $R_S = 0$ (which short circuits the input port), and $R_L = \infty$ (which open circuits the load port), (11.5) and (11.6) give

$$R_{\text{in}} = \frac{R_{\text{in}o}}{1 + PG_{io}\left(\infty, R_L\right)} \tag{11.11}$$

$$R_{\text{out}} = R_{\text{out}o}\left[1 + PG_{io}\left(R_S, 0\right)\right] \tag{11.12}$$

Diminished Closed-Loop Damping Factor

In addition to illuminating the driving-point and forward transfer characteristics of single-loop feedback architectures, the special case of global single-loop feedback illustrates the potenital instability problems pervasive of almost all feedback circuits. An examination of these problems begins by returning to (11.1) and letting the open-loop gain, G_o, be replaced by the two-pole frequency-domain function,

$$G_o\left(s\right) = \frac{G_o\left(0\right)}{\left(1 + \dfrac{s}{p_1}\right)\left(1 + \dfrac{s}{p_2}\right)} \tag{11.13}$$

where $G_o(0)$ symbolizes the zero-frequency open-loop gain. The pole frequencies p_1 and p_2 in (11.13) are either real numbers or complex conjugate pairs. Alternatively, (11.13) is expressible as

$$G_o(s) = \frac{G_o(0)}{1 + \dfrac{2\zeta_{\text{ol}}}{\omega_{\text{nol}}} s + \dfrac{s^2}{\omega_{\text{nol}}^2}} \tag{11.14}$$

where

$$\omega_{\text{nol}} = \sqrt{p_1 p_2} \tag{11.15}$$

represents the *undamped natural frequency* of oscillation of the open-loop configuration, and

$$\zeta_{\text{ol}} = \frac{1}{2}\left[\sqrt{\frac{p_2}{p_1}} + \sqrt{\frac{p_1}{p_2}}\right] \tag{11.16}$$

is the *damping factor* of the open-loop circuit.

In (11.1), let the feedback factor f be the single left-half-plane zero function,

$$f(s) = f_o\left(1 + \frac{s}{z}\right) \tag{11.17}$$

where z is the frequency of the real zero introduced by feedback, and f_{o} is the zero-frequency value of the feedback factor. The resultant loop gain is

$$T(s) = f_o\left(1 + \frac{s}{z}\right) G_o(s) \tag{11.18}$$

the zero-frequency value of the loop gain is

$$T(0) = f_o G_o(0) \tag{11.19}$$

and the zero frequency closed-loop gain $G_{\text{cl}}(0)$, is

$$G_{\text{cl}}(0) = \frac{G_o(0)}{1 + f_o G_o(0)} = \frac{G_o(0)}{1 + T(0)} \tag{11.20}$$

Upon inserting (11.14) and (11.17) into (11.1), the closed-loop transfer function is determined to be

$$G_{\text{cl}}(s) = \frac{G_{\text{cl}}(0)}{1 + \dfrac{2\zeta_{\text{cl}}}{\omega_{\text{ncl}}} s + \dfrac{s^2}{\omega_{ncl}^2}} \tag{11.21}$$

where the closed-loop undamped natural frequency of oscillation ω_{ncl} relates to its open-loop counterpart ω_{nol}, in accordance with

$$\omega_{\text{ncl}} = \omega_{\text{nol}} \sqrt{1 + T(0)} \tag{11.22}$$

Moreover, the closed-loop damping factor ζ_{cl} is

$$\zeta_{cl} = \frac{\zeta_{ol}}{\sqrt{1+T(0)}} + \left[\frac{T(0)}{1+T(0)}\right]\frac{\omega_{ncl}}{2z} = \frac{\zeta_{ol}}{\sqrt{1+T(0)}} + \left[\frac{T(0)}{\sqrt{1+T(0)}}\right]\frac{\omega_{nol}}{2z} \tag{11.23}$$

A frequency invariant feedback factor $f(s)$ applied to the open-loop configuration whose transfer function is given by (11.13) implies an infinitely large frequency, z, of the feedback zero. For this case, (11.23) confirms a closed-loop damping factor that is always less than the open-loop damping factor. Indeed, for a smaller than unity open-loop damping factor (which corresponds to complex conjugate open-loop poles) and reasonable values of the zero-frequency loop gain $T(0)$, $\zeta_{cl} \ll 1$. Thus, constant feedback applied around an underdamped two-pole open-loop amplifier yields a severely underdamped closed-loop configuration. It follows that the closed-loop circuit has a transient step response plagued by overshoot and a frequency response that displays response peaking within the closed-loop passband. Observe that underdamping is likely even in critically damped (identical real open-loop poles) or overdamped (distinct real poles) open-loop amplifiers, which, respectively, correspond to $\zeta_{ol} = 1$ and $\zeta_{ol} > 1$, when a large zero-frequency loop gain is exploited.

Underdamped closed-loop amplifiers are not unstable systems, but they are nonetheless unacceptable. From a practical design perspective, closed-loop underdamping predicted by relatively simple mathematical models of the loop gain portend undesirable amplifier responses or even closed-loop instability. The problem is that simple transfer function models invoked in a manual circuit analysis are oblivious to presumably second-order parasitic circuit layout and device model energy storage elements with effects that include a deterioration of phase and gain margins.

Frequency Invariant Feedback Factor

Let the open-loop amplifier be overdamped, such that its real satisfy the relationship

$$p_2 = \kappa^2 p_1 \tag{11.24}$$

If the open-loop amplifier pole p_1 is dominant, κ^2 is a real number that is greater than the magnitude, $|G_o(0)|$, of the open-loop zero frequency gain, which is presumed to be much larger than one. The open-loop damping factor in (11.16) resultantly reduces to $\zeta_{o1} \approx \kappa/2$. With $\kappa^2 > |G_o(0)| \gg 1$, which formally reflects the *dominant pole approximation*, the 3-dB bandwidth B_{ol} of the open-loop amplifier is given approximately by [15]

$$B_{ol} \approx \frac{\omega_{nol}}{2\zeta_{ol}} = \frac{1}{\dfrac{1}{p_1}+\dfrac{1}{p_2}} = \left(\frac{\kappa^2}{\kappa^2+1}\right)p_1 \tag{11.25}$$

As expected, (11.25) predicts an open-loop 3-dB bandwidth that is only slightly smaller than the frequency of the open-loop dominant pole.

The frequency, z, in (11.23) is infinitely large if frequency invariant degenerative feedback is applied around on open-loop amplifier. For a critically damped or overdamped closed-loop amplifier, $\zeta_{cl} > 1$. Assuming open-loop pole dominance, this constraint imposes the open-loop pole requirement,

$$\frac{p_2}{p_1} \geq 4\left[1+T(0)\right] \tag{11.26}$$

Thus, for large zero-frequency loop gain, $T(0)$, an underdamped closed-loop response is avoided if and only if the frequency of the nondominant open-loop pole is substantially larger than that of the dominant open-loop pole. Unless frequency compensation measures are exploited in the open loop, (11.26) is

difficult to satisfy, especially if feedback is implemented expressly to realize a substantive desensitization of response with respect to open-loop parameters. On the chance that (11.26) can be satisfied, and if the closed-loop amplifier emulates a dominant pole response, the closed-loop bandwidth is, using (11.22), (11.23), and (11.25),

$$B_{\mathrm{cl}} \approx \frac{\omega_{\mathrm{ncl}}}{2\zeta_{\mathrm{cl}}} \approx \left[1+T(0)\right]B_{\mathrm{ol}} \approx \left[1+T(0)\right]p_1 \tag{11.27}$$

Observe from (11.27) and (11.26) that the maximum possible closed-loop 3-dB bandwidth is 2 octaves below the minimum acceptable frequency of the nondominant open-loop pole.

Although (11.27) theoretically confirms the broadbanding property of negative feedback amplifiers, the attainment of very large closed-loop 3-dB bandwidths is nevertheless a challenging undertaking. The problem is that (11.26) is rarely satisfied. As a result, the open-loop configuration must be suitably compensated, usually by pole splitting methodology [16–18], to force the validity of (11.26). However, the open-loop poles are not mutually independent, so any compensation that increases p_2 is accompanied by decreases in p_1. The pragmatic upshot of the matter is that the closed-loop 3-dB bandwidth is not directly proportional to the uncompensated value of p_1 but instead, it is proportional to the smaller, compensated value of p_1.

Frequency Variant Feedback Factor (Compensation)

Consider now the case where the frequency, z, of the compensating feedback zero is finite and positive. Equation (11.23) underscores the stabilizing property of a left-half-plane feedback zero in that a sufficiently small positive z renders a closed-loop damping factor ζ_{cl} that can be made acceptably large, regardless of the value of the open-loop damping factor ζ_{ol}. To this end, $\zeta_{\mathrm{cl}} > 1/\sqrt{2}$ is a desirable design objective in that it ensures a monotonically decreasing closed-loop frequency response. If, as is usually a design goal, the open-loop amplifier subscribes to pole dominance, (11.23) translates the objective, $\zeta_{\mathrm{cl}} > 1/\sqrt{2}$, into the design constraint

$$z \le \frac{\left[\dfrac{T(0)}{1+T(0)}\right]\omega_{\mathrm{ncl}}}{\sqrt{2}-\dfrac{\omega_{\mathrm{ncl}}}{\left[1+T(0)\right]B_{\mathrm{ol}}}} \tag{11.28}$$

where use is made of (11.25) to cast ζ in terms of the open-loop bandwidth B_{ol}. When the closed-loop damping factor is precisely equal to $1/\sqrt{2}$ a maximally flat magnitude closed-loop response results for which the 3-dB bandwidth is ω_{ncl}. Equation (11.28) can then be cast into the more useful form

$$zG_{\mathrm{cl}}(0) = \frac{\mathrm{GBP}_{\mathrm{ol}}}{\sqrt{2}\left(\dfrac{\mathrm{GBP}_{\mathrm{ol}}}{\mathrm{GBP}_{\mathrm{cl}}}\right)-1} \tag{11.29}$$

where (11.20) is exploited, $\mathrm{GBP}_{\mathrm{ol}}$ is the *gain-bandwidth product* of the open-loop circuit, and $\mathrm{GBP}_{\mathrm{cl}}$ is the gain-bandwidth product of the resultant closed-loop network.

For a given open-loop gain-bandwidth product $\mathrm{GBP}_{\mathrm{ol}}$, a desired low-frequency closed-loop gain, $G_{\mathrm{cl}}(0)$, and a desired closed-loop gain-bandwidth product, $\mathrm{GBP}_{\mathrm{cl}}$, (11.29) provides a first-order estimate of the requisite feedback compensation zero. Additionally, note that (11.29) imposes an upper limit on the achievable high-frequency performance of the closed-loop configuration. In particular, because z must be positive to ensure acceptable closed-loop damping, (11.29) implies

$$\text{GBP}_{\text{ol}} > \frac{\text{GBP}_{\text{cl}}}{\sqrt{2}} \tag{11.30}$$

In effect, (11.30) imposes a lower limit on the required open-loop gain-bandwidth product commensurate with feedback compensation implemented to achieve a maximally flat, closed-loop frequency response.

11.5 Pole Splitting Open-Loop Compensation

Equation (11.26) underscores the desirability of achieving an open-loop dominant pole frequency response in the design of a feedback network. In particular, (11.26) shows that if the ultimate design goal is a closed-loop dominant pole frequency response, the frequency, p_2, of the nondominant open-loop amplifier pole must be substantially larger than its dominant pole counterpart, p_1. Even if closed-loop pole dominance is sacrificed as a trade-off for other performance merits, open-loop pole dominance is nonetheless a laudable design objective. This contention follows from (11.23) and (11.16), which combine to suggest that the larger p_2 is in comparison to p_1, the larger is the open-loop damping factor. In turn, the unacceptably underdamped closed-loop responses that are indicative of small, closed-loop damping factors are thereby eliminated. Moreover, (11.23) indicates that larger, open-loop damping factors impose progressively less demanding restrictions on the feedback compensation zero that may be required to achieve acceptable closed-loop damping. This observation is important because in an actual circuit design setting, small z in (11.23) generally translates into a requirement of a correspondingly large RC time constant, where implementation may prove difficult in monolithic circuit applications.

Unfortunately, many amplifiers, and particularly broadbanded amplifiers, earmarked for use as open-loop cells in degenerative feedback networks, are not characterized by dominant pole frequency responses. The frequency response of these amplifiers is therefore optimized in accordance with a standard design practice known as **pole splitting compensation**. Such compensation entails the connection of a small capacitor between two high impedance, phase inverting nodes of the open-loop topology [17, 19–21]. Pole splitting techniques increase the frequency p_2 of the uncompensated nondominant open-loop pole to a compensated value, say p_{2c}. The frequency, p_1, of the uncompensated dominant open-loop pole is simultaneously reduced to a smaller frequency, say p_{1c}. Although these pole frequency translations complement the design requirement implicit to (11.26) and (11.23), they do serve to limit the resultant closed-loop bandwidth, as discussed earlier. As highlighted next, they also impose other performance limitations on the open loop.

The Open-Loop Amplifier

The engineering methods, associated mathematics, and engineering trade-offs underlying pole splitting compensation are best revealed in terms of the generalized, phase inverting linear network abstracted in Figure 11.5. Although this amplifier may comprise the entire open-loop configuration, in the most general case, it is an interstage of the open loop. Accordingly, R_{st} in this diagram is viewed as the Thévenin equivalent resistance of either an input signal source or a preceding amplification stage. The response to the Thévenin driver, V_{st}, is the indicated output voltage, V_l, which is developed across the Thévenin load resistance, R_{lt}, seen by the stage under investigation. Note that the input current conducted by the amplifier is I_s, while the current flowing into the output port of the unit is denoted as I_{l}. The dashed branch containing the capacitor C_c, which is addressed later, is the pole splitting compensation element.

Because the amplifier under consideration is linear, any convenient set of two-port parameters can be used to model its terminal volt–ampere characteristics. Assuming the existence of the short circuit admittance, or y parameters,

$$\begin{bmatrix} I_s \\ I_l \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} \\ y_{21} & y_2 \end{bmatrix} \begin{bmatrix} V_i \\ V_l \end{bmatrix} \tag{11.31}$$

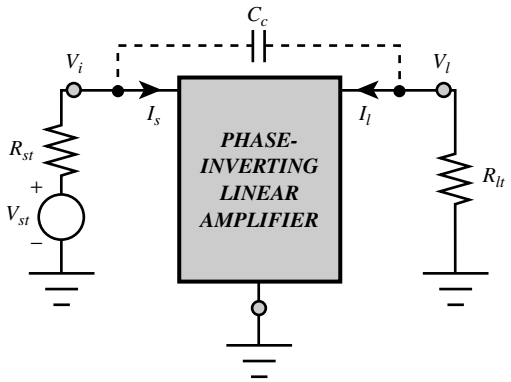


FIGURE 11.5 A linear amplifier for which a pole splitting compensation capacitance C_c is incorporated.

Defining

$$\begin{aligned} y_o &\triangleq y_{11} + y_{12} \\ y_o &\triangleq y_{22} + y_{12} \\ y_f &\triangleq y_{21} + y_{12} \\ y_r &\triangleq -y_{12} \end{aligned} \tag{11.32}$$

(11.31) implies

$$I_s = y_i V_i + y_r\left(V_i - V_l\right) \tag{11.33}$$

$$I_l = y_f V_i + y_o V_1 + y_r\left(V_l - V_i\right) \tag{11.34}$$

The last two expressions produce the *y*-parameter model depicted in Figure 7.6(a), in which y_i represents an effective shunt input admittance, y_o is a shunt output admittance, y_f is a forward transadmittance, and y_r reflects voltage feedback intrinsic to the amplifier.

Amplifiers amenable to pole splitting compensation have capacitive input and output admittances; that is, y_i and y_o are of the form

$$\begin{aligned} y_i &= \frac{1}{R_i} + sC_i \\ y_o &= \frac{1}{R_o} + sC_o \end{aligned} \tag{11.35}$$

Similarly,

$$\begin{aligned} y_f &= G_f - sC_f \\ y_r &= \frac{1}{R_r} + sC_r \end{aligned} \tag{11.36}$$

In (11.36), the conductance component G_f of the forward transadmittance y_f positive in a phase-inverting amplifier. Moreover, the reactive component $-sC_f$ of y_f produces an *excess phase angle*, and hence, a *group*

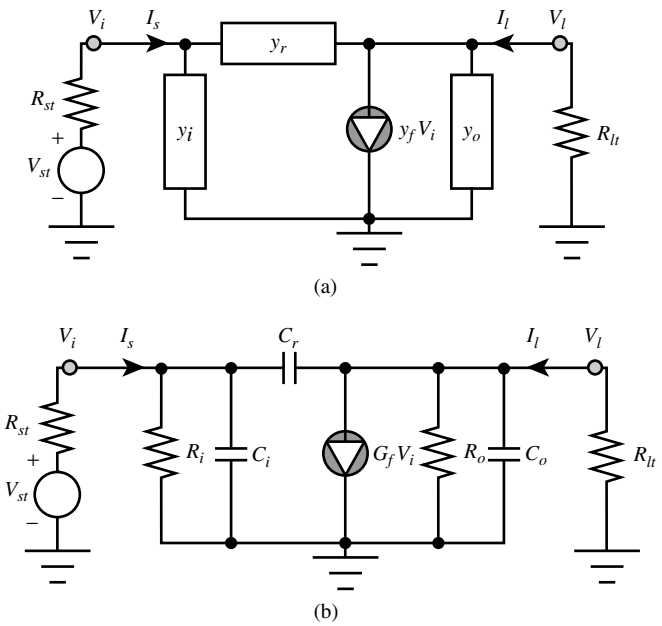


FIGURE 11.6 (a) The *y*-parameter equivalent circuit of the phase-inverting linear amplifier in Fig. 11.5. (b) An approximate form of the model in (a).

delay, in the forward gain function. This component, which deteriorates phase margin, can be ignored to first order if the signal frequencies of interest are not excessive in comparison to the upper-frequency limit of performance of the amplifier. Finally, the feedback internal to many practical amplifiers is predominantly capacitive so that the feedback resistance R_r can be ignored. These approximations allow the model in Figure 7.6(a) to be drawn in the form offered in Figure 11.6(b).

It is worthwhile interjecting that the six parameters indigenous to the model in Figure 11.6(b) need not be deduced analytically from the small-signal models of the active elements embedded in the subject interstage. Instead, SPICE can be exploited to evaluate the *y* parameters in (11.31) at the pertinent biasing level. Because these *y* parameters display dependencies on signal frequency, care should be exercised to evaluate their real and imaginary components in the neighborhood of the open-loop, 3-dB bandwidth to ensure acceptable computational accuracy at high frequencies. Once the *y* parameters in (11.31) are deduced by computer-aided analysis, the alternate admittance parameters in (11.23), as well as numerical estimates for the parameters, R_i, C_i, R_o, C_o, C_r, and G_f, in (11.35) and (11.36) follow straightforwardly.

Pole Splitting Analysis

An analysis of the circuit in Figure 11.6(b) produces a voltage transfer function $A_v(s)$ of the form

$$A_v(s) = \frac{V_l(s)}{V_{st}(s)} = A_v(0)\left[\frac{1-\dfrac{s}{z_r}}{\left(1+\dfrac{s}{p_1}\right)\left(1+\dfrac{s}{p_2}\right)}\right] \tag{11.37}$$

Letting

$$R_{ll} = R_{lt} \| R_o \tag{11.38}$$

an inspection of the circuit in Figure 11.6(b) confirms that

$$A_v(0) = -G_f R_{ll}\left(\frac{R_i}{R_i + R_{st}}\right) \tag{11.39}$$

is the zero frequency voltage gain. Moreover, the frequency, z_r, of the right-half-plane zero is

$$z_r = \frac{G_f}{C_r} \tag{11.40}$$

The lower pole frequency, p_1, and the higher pole frequency, p_2, derive implicitly from

$$\frac{1}{p_1} + \frac{1}{p_2} = R_{ll}(C_o + C_r) + R_{ss}\left[C_i + (1 + G_f R_{ll})C_r\right] \tag{11.41}$$

and

$$\frac{1}{p_1 p_2} = R_{ss} R_{ll} C_o\left[C_i + \left(\frac{C_o + C_i}{C_o}\right)C_r\right] \tag{11.42}$$

where

$$R_{ss} = R_{st} \triangleq R_i \tag{11.43}$$

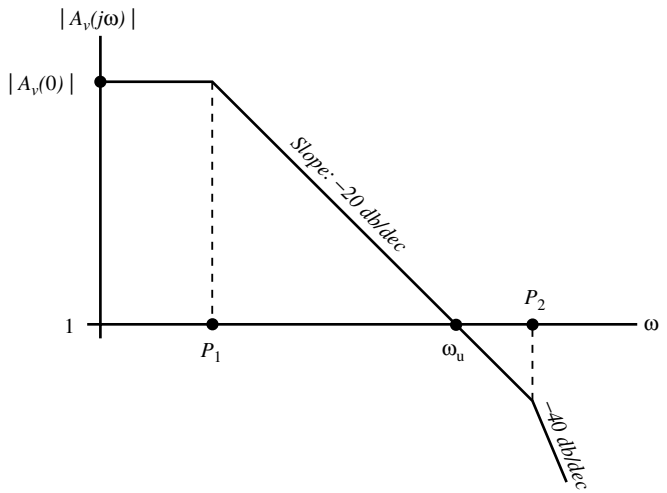
Most practical amplifiers, and particularly amplifiers realized in bipolar junction transistor technology, have very large forward transconductance, G_f, and small internal feedback capacitance, C_r. The combination of large G_f and small C_r renders the frequency in (11.40) so large as to be inconsequential to the passband of interest. When utilized in a high-gain application, such as the open-loop signal path of a feedback amplifier, these amplifiers also operate with a large effective load resistance, R_{ll}. Accordingly, (11.41) can be used to approximate the pole frequency p_1 as

$$p_1 \approx \frac{1}{R_{ss}\left[C_i + (1 + G_f R_{ll})C_r\right]} \tag{11.44}$$

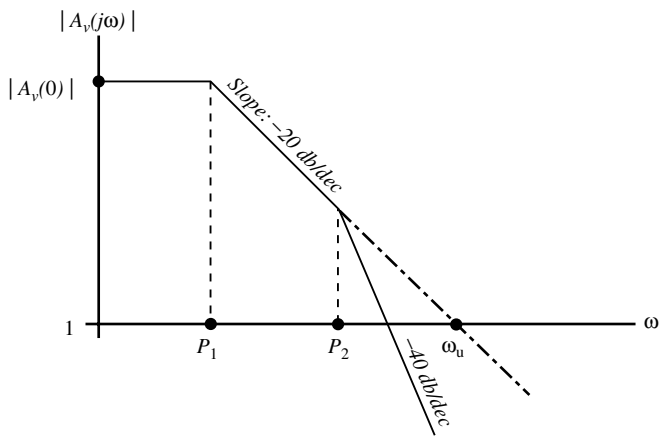
Substituting this result into (11.42), the approximate frequency p_2 of the high-frequency pole is

$$p_2 \approx \frac{C_i + (1 + G_f R_{ll})C_r}{R_{ll} C_o\left[C_i + \left(\frac{C_o + C_i}{C_o}\right)C_r\right]} \tag{11.45}$$

Figure 11.7 illustrates asymptotic frequency responses corresponding to pole dominance and to a two-pole response. Figure 11.7(a) depicts the frequency response of a dominant pole amplifier, which does not require pole splitting compensation. Observe that its high-frequency response is determined by a single pole (p_1 in this case) through the signal frequency at which the gain ultimately degrades to unity. In this interpretation of a dominant pole amplifier, p_2 is not only much larger than p_1, but is in fact larger than the unity gain frequency, which is indicated as ω_u in the figure. This unity gain frequency, which can be viewed as an upper limit to the useful passband of the amplifier, is approximately, $|A_v(0)|p_1$. To the extent that p_1 is essentially the 3-dB bandwidth when $p_2 \gg p_1$, the unity gain frequency is also the



(a)



(b)

FIGURE 11.7 (a) Asymptotic frequency response for a dominant pole amplifier. Such an amplifier does not require pole splitting compensation because the two lowest frequency amplifier poles, p_1 and p_2, are already widely separated. (b) The frequency response of an amplifier with high-frequency response that is strongly influenced by both of its lowest frequency poles. The basic objective of pole splitting compensation is to transform the indicated frequency response to a form that emulates that depicted in (a).

gain-bandwidth product (GBP) of the subject amplifier. In short, with $|A_v(j\omega_u)| \triangleq 1$, $p_2 \gg p_1$ in (11.37) implies

$$\omega_u \approx |A_v(0)|p_1 \approx \text{GBP} \tag{11.46}$$

The contrasting situation of a response indigenous to the presence of two significant open-loop poles is illustrated in Figure 11.7(b). In this case, the higher pole frequency p_2 is smaller than ω_u and hence, the amplifier does not emulate a single-pole response throughout its theoretically useful frequency range. The two critical frequencies, p_1 and p_2, remain real numbers, and as long as $p_2 \neq p_1$, the corresponding damping factor, is greater than one. However, the damping factor of the two-pole amplifier (its response is plotted in Figure 11.7(b)) is nonetheless smaller than that of the dominant pole amplifier. It follows that, for reasonable loop gains, unacceptable underdamping is more likely when feedback is invoked around the

two-pole amplifier, as opposed to the same amount of feedback applied around a dominant pole amplifier. Pole splitting attempts to circumvent this problem by transforming the pole conglomeration of the two pole amplifier into one that emulates the dominant pole situation inferred by Figure 11.7(a).

To the foregoing end, append the compensation capacitance C_c between the input and the output ports of the phase-inverting linear amplifier, as suggested in Figure 11.5. With reference to the equivalent circuit in Figure 11.6(b), the electrical impact of this additional element is the effective replacement of the internal feedback capacitance C_r by the capacitance sum $(C_r + C_c)$. Letting

$$C_p \triangleq C_r + C_c \tag{11.47}$$

it is apparent that (11.40)–(11.42) remain applicable, provided that C_r in these relationships is supplanted by C_p. Because C_p is conceivably significantly larger than C_c, however, the approximate expressions for the resultant pole locations differ from those of (11.44) and (11.45). In particular, a reasonable approximation for the compensated value, say P_{1c}, of the lower pole frequency is now

$$p_{1c} \approx \frac{1}{\left[R_{ll} + \left(1 + G_f R_{ll}\right) R_{ss}\right] C_p} \tag{11.48}$$

while the higher pole frequency, p_{2c}, becomes

$$p_{2c} \approx \frac{1}{\left(R_{ss} \| R_{ll} \| \dfrac{1}{G_f}\right)\left(C_o + C_i\right)} \tag{11.49}$$

Clearly, $p_{1c} < p_1$ and $p_{2c} > p_2$. Moreover, for large G_f, p_{2c} is potentially much larger than p_{1c}. It should also be noted that the compensated value, say, z_{rc}, of the right-half-plane zero is smaller than its uncompensated value, z_r, because (11.40) demonstrates that

$$z_{rc} = \frac{G_f}{C_p} = z_r\left(\frac{C_r}{C_r + C_c}\right) \tag{11.50}$$

Although z_{rc} can conceivably exert a significant influence on the high-frequency response of the compensated amplifier, the following discussion presumes tacitly that $z_{rc} > p_{2c}$ [2].

Assuming a dominant pole frequency response, the compensated unity gain frequency, ω_{uc}, is, using (11.39), (11.46), and (11.48),

$$\omega_{uc} \approx \left|A_v(0)\right| p_{1c} \approx \left(\frac{1}{R_{st} C_p}\right)\left[G_f\left(R_{ss} \| R_{ll} \| \frac{1}{G_f}\right)\right] \tag{11.51}$$

It is interesting to note that

$$\omega_{uc} < \left(\frac{1}{R_{st} C_p}\right) \tag{11.52}$$

that is, the unity gain frequency is limited by the inverse of the RC time constant formed by the Thévenin source resistance R_{st} and the net capacitance Cp appearing between the input port and the phase inverted output port. The subject inequality comprises a significant performance limitation, for if p_{2c} is indeed

much larger than p_{ic}, ω_{uc} is approximately the GBP of the compensated cell. Accordingly, for a given source resistance, a required open-loop gain, and a desired open-loop bandwidth, (11.52) imposes an upper limit on the compensation capacitance that can be exploited for pole splitting purposes.

In order for the compensated amplifier to behave as a dominant pole configuration, p_{2c} must exceed ω_{uc}, as defined by (11.51). Recalling (11.49), the requisite constraint is found to be

$$R_{\mathrm{st}}C_p > G_f\left(R_{ss}\|R_{ll}\|\frac{1}{G_f}\right)^2\left(C_o + C_i\right) \tag{11.53}$$

Assuming $G_f(R_{ss}/R_{ll}) \ll 1$, (11.53) reduces to the useful simple form

$$C_f R_{\mathrm{st}} > \frac{C_o + C_i}{C_p} \tag{11.54}$$

which confirms the need for large forward transconductance G_f if pole splitting is to be an effective compensation technique.

11.6 Summary

The use of negative feedback is fundamental to the design of reliable and reproducible analog electronic networks. Accordingly, this chapter documents the salient features of the theory that underlies the efficient analysis and design of commonly used feedback networks. Four especially significant points are postulated in this section.

1. By judiciously exploiting signal flow theory, the classical expression, (11.1), for the I/O transfer relationship of a linear feedback system is rendered applicable to a broad range of electronic feedback circuits. This expression is convenient for design-oriented analysis because it clearly identifies the open-loop gain, G_o, and the loop gain, T. The successful application of signal flow theory is predicated on the requirement that the feedback factor, to which T is proportional and that appears in the signal flow literature as a "critical" or "reference" parameter, can be identified in a given feedback circuit.
2. Signal flow theory, as applied to electronic feedback architectures, proves to be an especially expedient analytical tool because once the loop gain T is identified, the driving-point input and output impedances follow with minimal additional calculations. Moreover, the functional dependence of T on the Thévenin source and terminating load impedances unambiguously brackets the magnitudes of the driving point I/O impedances attainable in particular types of feedback arrangements.
3. The damping factor concept is advanced herewith as a simple way of assessing the relative stability of both the open and closed loops of a feedback circuit. The open-loop damping factor derives directly from the critical frequencies of the open-loop gain, while these frequencies and any zeros appearing in the loop gain unambiguously define the corresponding closed-loop damping factor. Signal flow theory is once again used to confirm the propensity of closed loops toward instability unless the open-loop subcircuit functions as a dominant pole network. Also confirmed is the propriety of the common practice of implementing a feedback zero as a means of stabilizing an otherwise potentially unstable closed loop.
4. Pole splitting as a means to achieve dominant pole open-loop responses is definitively discussed. Generalized design criteria are formulated for this compensation scheme, and limits of performance are established. Of particular interest is the fact that pole splitting limits the GBP of the compensated amplifier to a value that is determined by a source resistance-compensation capacitance time constant.

References

[1] J. A. Mataya, G. W. Haines, and S. B. Marshall, "IF amplifier using C_c -compensated transistors," *IEEE J. Solid-State Circuits*, vol. SC-3, pp. 401–407, Dec. 1968.

[2] W. G. Beall and J. Choma, Jr., "Charge-neutralized differential amplifiers," *J. Analog Integrat. Circuits Signal Process.*, vol. 1, pp. 33–44, Sep. 1991.

[3] J. Choma, Jr., "A generalized bandwidth estimation theory for feedback amplifiers," *IEEE Trans. Circuits Syst.*, vol. CAS-31, pp. 861–865, Oct. 1984.

[4] R. D. Thornton, C. L. Searle, D. O. Pederson, R. B. Adler, and E. J. Angelo, Jr., *Multistage Transistor Circuits*, New York: John Wiley & Sons, 1965, chaps. 1, 8.

[5] H. W. Bode, *Network Analysis and Feedback Amplifier Design*, New York: Van Nostrand, 1945.

[6] P. J. Hurst, "A comparison of two approaches to feedback circuit analysis," *IEEE Trans Education*, vol. 35, pp. 253–261, Aug. 1992.

[7] M. S. Ghausi, *Principles and Design of Linear Active Networks*, New York: McGraw-Hill, 1965, pp. 40–56.

[8] A. J. Cote, Jr. and J. B. Oakes, *Linear Vacuum-Tube and Transistor Circuits*, New York: McGraw-Hill, 1961, pp. 40–46.

[9] S. J. Mason, "Feedback theory — Some properties of signal flow graphs," *Proc. IRE*, vol. 41, pp. 1144–1156, Sep. 1953.

[10] S. J. Mason, "Feedback theory — Further properties of signal flow graphs," *Proc. IRE*, vol. 44, pp. 920–926, July 1956.

[11] N. Balabanian and T. A. Bickart, *Electrical Network Theory*, New York: John Wiley & Sons, 1969, pp. 639–669.

[12] J. Choma, Jr., "Signal flow analysis of feedback networks," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 455–463, April 1990.

[13] J. Choma, Jr., *Electrical Networks: Theory and Analysis*, New York: Wiley Interscience, 1985, pp. 589–605.

[14] P. J. Hurst, "Exact simulation of feedback circuit parameters," *IEEE Trans. Circuits Syst.*, vol. 38, pp. 1382–1389, Nov. 1991.

[15] J. Choma, Jr. and S. A. Witherspoon, "Computationally efficient estimation of frequency response and driving point impedance in wideband analog amplifiers," *IEEE Trans. Circuits Syst.*, vol. 37, pp. 720–728, June 1990.

[16] R. G. Meyer and R. A. Blauschild, "A wide-band low-noise monolithic transimpedance amplifier," *IEEE J. Solid-State Circuits*, vol. SC-21, pp. 530–533, Aug. 1986.

[17] Y. P. Tsividis, "Design considerations in single-channel MOS analog integrated circuits," *IEEE J. Solid-State Circuits*, vol. SC-13, pp. 383–391, June 1978.

[18] J. J. D'Azzo and C. H. Houpis, *Feedback Control System Analysis and Synthesis*, New York: McGraw-Hill, 1960, pp. 230–234.

[19] P. R. Gray and R. G. Meyer, *Analysis and Design of Analog Integrated Circuits*, New York: John Wiley & Sons, 1977, pp. 512–521.

[20] P. R. Gray, "Basic MOS operational amplifier design — An overview," in *Analog MOS Integrated Circuits*, P. R. Gray, D. A. Hodges, and R. W. Brodersen, Eds., New York: IEEE, 1980, pp. 28–49.

[21] J. E. Solomon, "The monolithic op-amp: A tutorial study," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 314–332, Dec. 1974.

12

Feedback Amplifier Configurations

John Choma, Jr.
University of Southern California

12.1 Introduction

Four basic types of single-loop feedback amplifiers are available: the **series-shunt**, **shunt-series**, **shunt-shunt**, and **series-series architectures** [1]. Each of these cells is capable of a significant reduction of the dependence of forward transfer characteristics on the ill-defined or ill-controlled parameters implicit to the open-loop gain; but none of these architectures can simultaneously offer controlled driving-point input and output impedances. Such additional control is afforded only by dual global loops comprised of series and/or shunt feedback signal paths appended to an open-loop amplifier [2], [3]. Only two types of global dual-loop feedback architectures are used: the **series-series/shunt-shunt feedback amplifier** and the **series-shunt/shunt-series feedback amplifier**.

Although only bipolar technology is exploited in the analysis of the aforementioned four single-loop and two dual-loop feedback cells, all disclosures are generally applicable to metaloxide-silicon (MOS), heterostructure bipolar transistor (HBT), and III–V compound metal-semiconductor field-effect transistor (MESFET) technologies. All analytical results derive from an application of a hybrid, signal flow/two-port parameter analytical tack. Because the thought processes underlying this technical approach apply to all feedback circuits, the subject analytical procedure is developed in detail for only the series-shunt feedback amplifier.

12.2 Series-Shunt Feedback Amplifier

Circuit Modeling and Analysis

Figure 12.1(a) depicts the **ac schematic diagram** (a circuit diagram divorced of biasing details) of a series-shunt feedback amplifier. In this circuit, the output voltage V_O, which is established in response to a single source represented by the Thévenin voltage V_{ST}, and the Thévenin resistance, R_{ST}, is sampled by

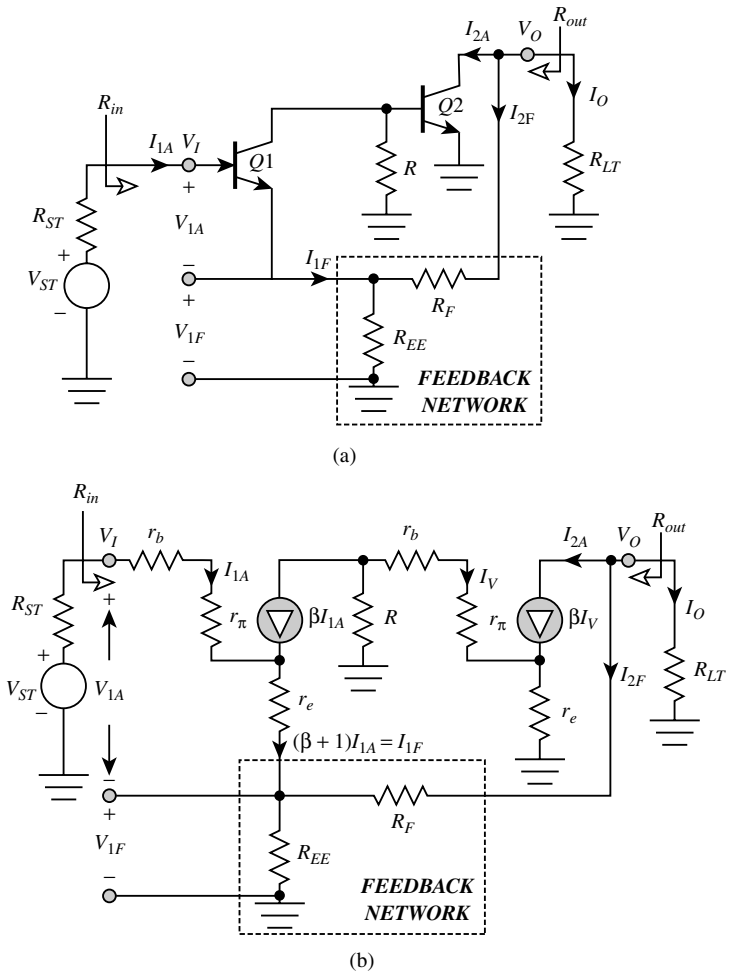


FIGURE 12.1 (a) The ac schematic diagram of a bipolar series-shunt feedback amplifier. (b) Low-frequency small-signal equivalent circuit of the feedback amplifier.

the feedback network composed of the resistances, R_{EE} and R_F. The sampled voltage is fed back in such a way that the closed-loop input voltage, V_I, is the sum of the voltage, V_{1A}, across the input port of the amplifier and the voltage V_{1F}, developed across R_{EE} in the feedback subcircuit. Because $V_I = V_{1A} + V_{1F}$, the output port of the feedback configuration can be viewed as connected in series with the amplifier input port. On the other hand, output voltage sampling constrains the net load current, I_O, to be the algebraic sum of the amplifier output port current, I_{2A}, and the feedback network input current, I_{2F}. Accordingly, the output topology is indicative of a shunt connection between the feedback subcircuit and the amplifier output port. The fact that voltage is fed back to a voltage-driven input port renders the driving point input resistance, R_{in}, of the closed-loop amplifier large, whereas the driving-point output resistance, R_{out}, seen by the terminating load resistance, R_{LT}, is small. The resultant closed-loop amplifier is therefore best suited for voltage amplification, in the sense that the closed-loop voltage gain, V_O/V_{ST}, can be made approximately independent of source and load resistances. For large loop gain, this voltage transfer function is also nominally independent of transistor parameters.

Assuming that transistors $Q1$ and $Q2$ are identical devices that are biased identically, Figure 12.1(b) is the applicable low-frequency equivalent circuit. This equivalent circuit exploits the hybrid-π model [4] of a bipolar junction transistor, subject to the proviso that the forward Early resistance [5] used to emulate base conductivity modulation is sufficiently large to warrant its neglect. Because an infinitely

large forward Early resistance places the internal collector resistance (not shown in the figure) of a bipolar junction transistor in series with the current controlled current source, this collector resistance can be ignored as well.

The equivalent circuit of Figure 12.1(b) can be reduced to a manageable topology by noting that the ratio of the signal current, I_V, flowing into the base of transistor $Q2$ to the signal current, I_{1A}, flowing into the base of transistor $Q1$ is

$$\frac{I_V}{I_{1A}} \triangleq -K_\beta = -\frac{\beta R}{R + r_b + r_\pi + (\beta + 1) r_e} = -\frac{\alpha R}{r_{ib} + (1 - \alpha) R} \tag{12.1}$$

where

$$\alpha = \frac{\beta}{\beta + 1} \tag{12.2}$$

is the small-signal, short-circuit common base current gain, and

$$r_{ib} = r_e + \frac{r_\pi + r_b}{\beta + 1} \tag{12.3}$$

symbolizes the short-circuit input resistance of a common base amplifier. It follows that the current source βI_v in Figure 12.1(b) can be replaced by the equivalent current $(-\beta K_\beta I_{1A})$.

A second reduction of the equivalent circuit in Figure 12.1(b) results when the feedback subcircuit is replaced by a model that reflects the *h*-parameter relationships

$$\begin{bmatrix} V_{1F} \\ I_{2F} \end{bmatrix} = \begin{bmatrix} h_{if} & h_{rf} \\ h_{ff} & h_{of} \end{bmatrix} \begin{bmatrix} I_{1F} \\ V_o \end{bmatrix} \tag{12.4}$$

where $V_{1F}(V_O)$ represents the signal voltage developed across the output (input) port of the feedback subcircuit and $I_{1F}(I_{2F})$ symbolizes the corresponding current flowing into the feedback output (input) port. Although any homogeneous set of two-port parameters can be used to model the feedback subcircuit, *h* parameters are the most convenient selection herewith. In particular, the feedback amplifier undergoing study is a series-shunt configuration. The *h*-parameter equivalent circuit represents its input port as a Thévenin circuit and its input port as a Norton configuration, therefore, the *h*-parameter equivalent circuit is likewise a series-shunt structure.

For the feedback network at hand, which is redrawn for convenience in Figure 12.2(a), the *h*-parameter equivalent circuit is as depicted in Figure 12.2(b). The latter diagram exploits the facts that the short-circuit input resistance h_{if} is a parallel combination of the resistance R_{EE} and R_F, and the open-circuit output conductance h_{of}, is $1/(R_{\text{EE}} + R_F)$. The open-circuit reverse voltage gain h_{rf} is

$$h_{rf} = \frac{R_{\text{EE}}}{R_{\text{EE}} + R_F} \tag{12.5}$$

while the short-circuit forward current gain h_{ff} is

$$h_{ff} = \frac{R_{\text{EE}}}{R_{\text{EE}} + R_F} = -h_{rf} \tag{12.6}$$

Figure 12.2(c) modifies the equivalent circuit in Figure 12.2(b) in accordance with the following two arguments. First, h_{rf} in (12.5) is recognized as the fraction of the feedback subcircuit input signal that is

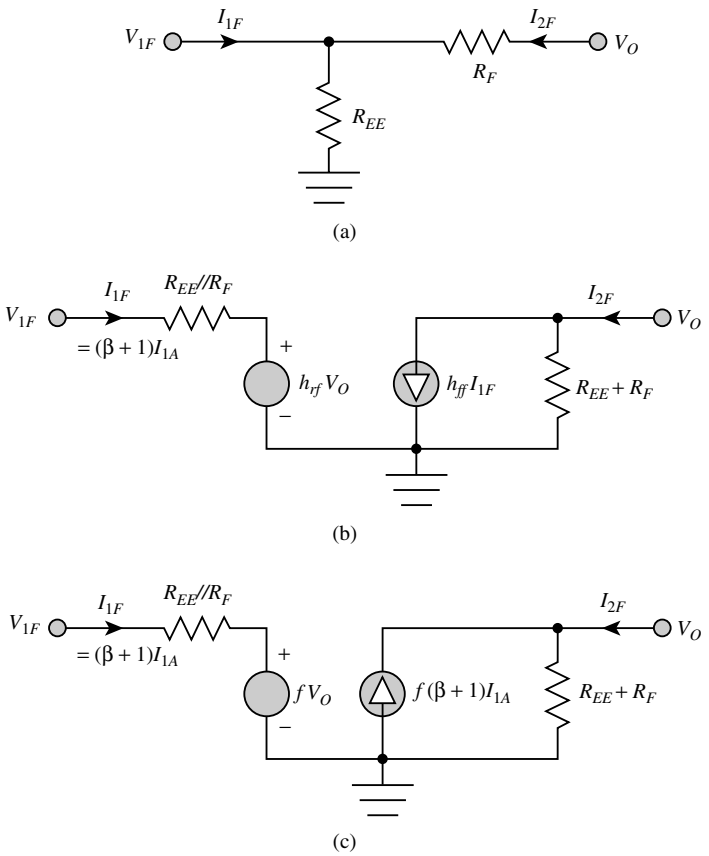


FIGURE 12.2 (a) The feedback subcircuit in the series-shunt feedback amplifier of Figure 12.1(a). (b) The *h*-parameter equivalent circuit of the feedback subcircuit. (c) Alternative form of the *h*-parameter equivalent circuit.

fed back as a component of the feedback subcircuit output voltage, V_{1F}. But this subcircuit input voltage is identical to the closed-loop amplifier output signal V_O. Moreover, V_{1F} superimposes with the Thévenin input signal applied to the feedback amplifier to establish the amplifier input port voltage, V_{1A}. It follows that h_{rf} is logically referenced as a feedback factor, say f, of the amplifier under consideration; that is,

$$h_{rf} = \frac{R_{\text{EE}}}{R_{\text{EE}} + R_F} \triangleq f \tag{12.7}$$

and by (12.6),

$$h_{ff} = -\frac{R_{\text{EE}}}{R_{\text{EE}} + R_F} = -f \tag{12.8}$$

Second, the feedback subcircuit output current, I_{1F}, is, as indicated in Figure 12.1(b), the signal current, $(\beta + 1)I_{1A}$. Thus, in the model of Figure 12.2(b),

$$h_{ff}I_{1F} = -f(\beta + 1)I_{1A} \tag{12.9}$$

If the model in Figure 12.2(c) is used to replace the feedback network in Figure 12.1(b) the equivalent circuit of the series-shunt feedback amplifier becomes the alternative structure offered in Figure 12.3. In

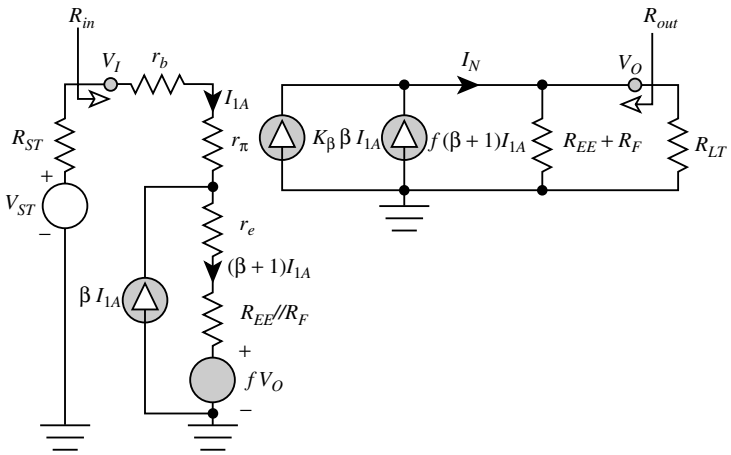


FIGURE 12.3 Modified small-signal model of the series-shunt feedback amplifier.

arriving at this model, care has been exercised to ensure that the current flowing through the emitter of transistor $Q1$ is $(\beta + 1)I_{1A}$. It is important to note that the modified equivalent circuit delivers transfer and driving point impedance characteristics that are identical to those implicit to the equivalent circuit of Figure 12.1(b). In particular, the traditional analytical approach to analyzing a series-shunt feedback amplifier tacitly presumes the satisfaction of the Brune condition [6] to formulate a composite structure where the *h*-parameter matrix is the sum of the respective *h*-parameter matrices for the open-loop and feedback circuits. In contrast, the model of Figure 12.3 derives from Figure 12.1(b) without invoking the Brune requirement, which is often not satisfied. It merely exploits the substitution theorem; that is, the feedback network in Figure 12.1(b) is substituted by its *h*-parameter representation.

In addition to modeling accuracy, the equivalent circuit in Figure 12.3 boasts at least three other advantages. The first is an illumination of the vehicle by which feedback is implemented in the series-shunt configuration. This vehicle is the voltage controlled voltage source, $f\,V_O$, which feeds back a fraction of the output signal to produce a branch voltage that algebraically superimposes with, and thus modifies, the applied source voltage effectively seen by the input port of the open-loop amplifier. Thus, with $f = 0$, no feedback is evidenced, and the model at hand emulates an open-loop configuration. But even with $f = 0$, the transfer and driving-point impedance characteristics of the resultant open-loop circuit are functionally dependent on the feedback elements, R_{EE} and R_F, because appending the feedback network to the open-loop amplifier incurs additional impedance loads at both the input and the output ports of the amplifier.

The second advantage of the subject model is its revelation of the magnitude and nature of feed-forward through the closed loop. In particular, note that the signal current, I_N, driven into the effective load resistance comprised of the parallel combination of $(R_{EE} + R_F)$ and R_{LT}, is the sum of two current components. One of these currents, $\beta K_\beta I_{1A}$, materializes from the transfer properties of the two transistors utilized in the amplifier. The other current, $f\,(\beta + 1)I_{1A}$, is the feed-forward current resulting from the bilateral nature of the passive feedback network. In general, negligible feed-forward through the feedback subcircuit is advantageous, particularly in high-frequency signal-processing applications. To this end, the model in Figure 12.3 suggests the design requirement,

$$f \ll \alpha K_\beta \tag{12.10}$$

When the resistance, R, in Figure 12.1(a) is the resistance associated with the output port of a PNP current source used to supply biasing current to the collector of transistor $Q1$ and the base of transistor $Q2$, K_β approaches β, and (12.10) is easily satisfied; however, PNP current sources are undesirable in broadband low-noise amplifiers. In these applications, the requisite biasing current must be supplied by

a passive resistance, R, connected between the positive supply voltage and the junction of the $Q1$ collector and the $Q2$ base. Unfortunately, the corresponding value of K_β can be considerably smaller than β, with the result that (12.10) may be difficult to satisfy. Circumvention schemes for this situation are addressed later.

A third attribute of the model in Figure 12.3 is its disposition to an application of signal flow theory. For example, with the feedback factor f selected as the reference parameter for signal flow analysis, the open-loop voltage gain $G_{vo}(R_{ST}, R_{LT})$, of the series-shunt feedback amplifier is computed by setting f to zero. Assuming that (12.10) is satisfied, circuit analysis reveals this gain as

$$G_{vo}\left(R_{ST}, R_{LT}\right) = \alpha K_\beta \left[\frac{\left(R_{EE} + R_F\right) \| R_{LT}}{r_{ib} + \left(1-\alpha\right) R_{ST} + \left(R_{EE} \| R_F\right)} \right] \tag{12.11}$$

The corresponding input and output driving point resistances, R_{ino} and R_{outo}, respectively, are

$$R_{ino} = r_B + r_\pi + \left(\beta + 1\right)\left(r_E + R_{EE} \| R_F\right) \tag{12.12}$$

and

$$R_{outo} = R_{EE} + R_F \tag{12.13}$$

It follows that the closed-loop gain $G_v\left(R_{ST}, R_{LT}\right)$ of the series-shunt feedback amplifier is

$$G_v\left(R_{ST}, R_{LT}\right) = \frac{G_{vo}\left(R_{ST}, R_{LT}\right)}{1 + T} \tag{12.14}$$

where the loop gain T is

$$\begin{aligned} T = f G_{vo}\left(R_{ST}, R_{LT}\right) &= \left(\frac{R_{EE}}{R_{EE} + R_F} \right) G_{vo}\left(R_{ST}, R_{LT}\right) \\ &= \alpha K_\beta \left(\frac{R_{EE}}{R_{EE} + R_F + R_{LT}} \right) \left[\frac{R_{LT}}{r_{ib} + \left(1-\alpha\right) R_{ST} + \left(R_{EE} \| R_F\right)} \right] \end{aligned} \tag{12.15}$$

For $T \gg 1$, which mandates a sufficiently large K_β in (12.11), the closed-loop gain collapses to

$$G_v\left(R_{ST}, R_{LT}\right) \approx \frac{1}{f} = 1 + \frac{R_F}{R_{EE}} \tag{12.16}$$

which is independent of active element parameters. Moreover, to the extent that $T \gg 1$ the series-shunt feedback amplifier behaves as an ideal voltage controlled voltage source in the sense that its closed-loop voltage gain is independent of source and load terminations. The fact that the series-shut feedback network behaves approximately as an ideal voltage amplifier implies that its closed-loop driving point input resistance is very large and its closed-loop driving point output resistance is very small. These facts are confirmed analytically by noting that

$$\begin{aligned} R_{in} = R_{ino}\left[1 + f G_{vo}\left(0, R_L\right)\right] &\approx f R_{ino} G_{vo}\left(0, R_L\right) \\ &= \beta K_\beta \left(\frac{R_{EE}}{R_{EE} + R_F + R_{LT}} \right) R_{LT} \end{aligned} \tag{12.17}$$

and

$$R_{\text{out}} = \frac{R_{\text{outo}}}{1 + f G_{vo}\left(R_S, \infty\right)} \approx \frac{R_{\text{outo}}}{f G_{vo}\left(R_S, \infty\right)} = \left(1 + \frac{R_F}{R_{\text{EE}}}\right)\left[\frac{r_{ib} + (1-\alpha)R_{\text{ST}} + R_{\text{EE}} \| R_F}{\alpha K_\beta}\right] \tag{12.18}$$

To the extent that the interstage biasing resistance, R, is sufficiently large to allow K_β to approach β, observe that R_{in} in (12.17) is nominally proportional to β^2, while R_{out} in (12.18) is inversely proportional to β.

Feed-Forward Compensation

When practical design restrictions render the satisfaction of (12.10) difficult, feed-forward problems can be circumvented by inserting an emitter follower between the output port of transistor $Q2$ in the circuit diagram of Figure 12.1(a) and the node to which the load termination and the input terminal of the feedback subcircuit are incident [2]. The resultant circuit diagram, inclusive now of simple biasing subcircuits, is shown in Figure 12.4. The buffer transistor $Q3$ increases the original short-circuit forward current gain, $K_\beta\beta$, of the open-loop amplifier by a factor approaching $(\beta + 1)$, while not altering the feed-forward factor implied by the feedback network in Figure 12.1(a). In effect, K_β is increased by a factor of almost $(\beta + 1)$, thereby making (12.10) easy to satisfy. Because of the inherently low output resistance of an emitter follower, the buffer also reduces the driving-point output resistance achievable by the original configuration.

The foregoing contentions can be confirmed through an analysis of the small-signal model for the modified amplifier in Figure 12.4. Such an analysis is expedited by noting that the circuit to the left of the current controlled current source, $K_\beta\beta I_{1A}$, in Figure 12.3 remains applicable. For zero feedback, it follows that the small-signal current I_{1A} flowing into the base of transistor $Q1$ derives from

$$\left.\frac{I_{1A}}{V_{\text{ST}}}\right|_{f=0} = \frac{1-\alpha}{r_{ib} + (1-\alpha)R_{\text{ST}} + \left(R_{\text{EE}} \| R_F\right)} \tag{12.19}$$

The pertinent small-signal model for the buffered series-shunt feedback amplifier is resultantly the configuration offered in Figure 12.5.

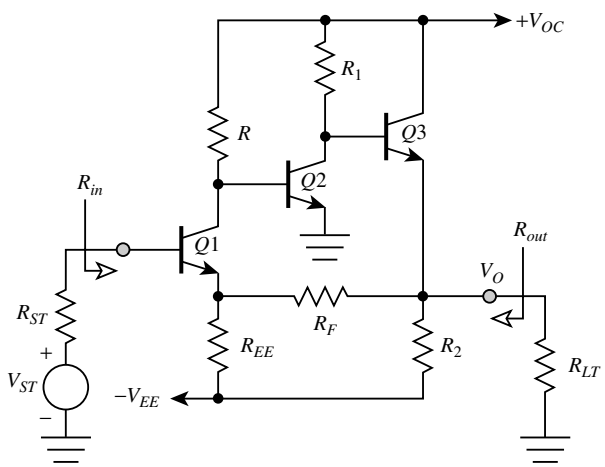


FIGURE 12.4 A series-shunt feedback amplifier that incorporates an emitter follower output stage to reduce the effects of feed-forward through the feedback network.

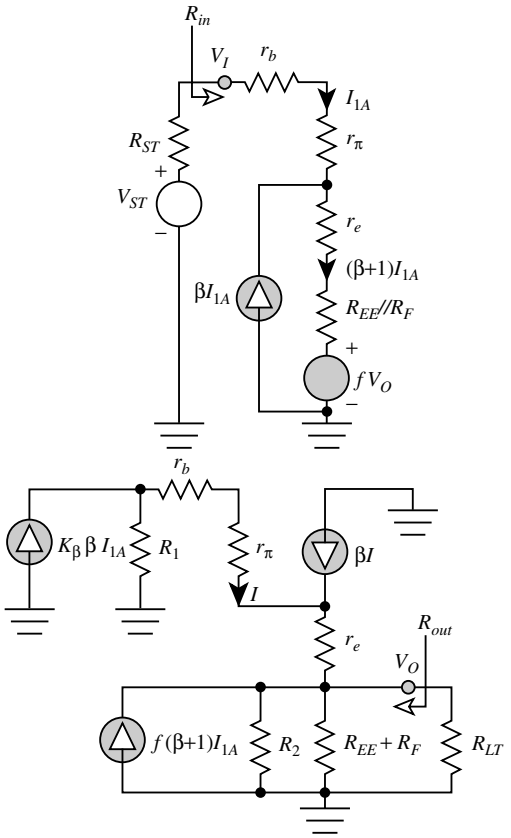


FIGURE 12.5 Small-signal model of the buffered series-shunt feedback amplifier.

Letting

$$R' = R_2 \left\| \left(R_{\mathrm{EE}} + R_F\right) \right\| R_{\mathrm{LT}} \tag{12.20}$$

an analysis of the structure in Figure 12.5 reveals

$$\frac{V_O}{I_{1A}} = (\beta+1)\left[\frac{R'}{R' + r_{ib} + (1-\alpha)R_1}\right]\left\{\alpha K_\beta R_1 + f\left[r_{ib} + (1-\alpha)R_1\right]\right\} \tag{12.21}$$

which suggests negligible feed-forward for

$$f \ll \frac{\alpha K_\beta R_1}{r_{ib} + (1-\alpha)R_1} \tag{12.22}$$

Note that for large R_1, (12.22) implies the requirement $f \ll \beta K_\beta$, which is easier to satisfy than is (12.10). Assuming the validity of (12.22), (12.21), and (12.19) deliver an open-loop voltage gain,. $G_{vo}(R_{\mathrm{ST}}, R_{\mathrm{LT}})$, of

$$G_{vo}\left(R_{\mathrm{ST}}, R_{\mathrm{LT}}\right) = \alpha K_\beta \left[\frac{R'}{r_{ib} + (1-\alpha)R_{\mathrm{ST}} + R_{\mathrm{EE}} \| R_F}\right]\left[\frac{R_1}{R' + r_{ib} + (1-\alpha)R_1}\right] \tag{12.23}$$

Recalling (12.1), which demonstrates that K_β approaches β for large R, (12.23) suggests an open-loop gain that is nominally proportional to β^2 if R_1 is also large.

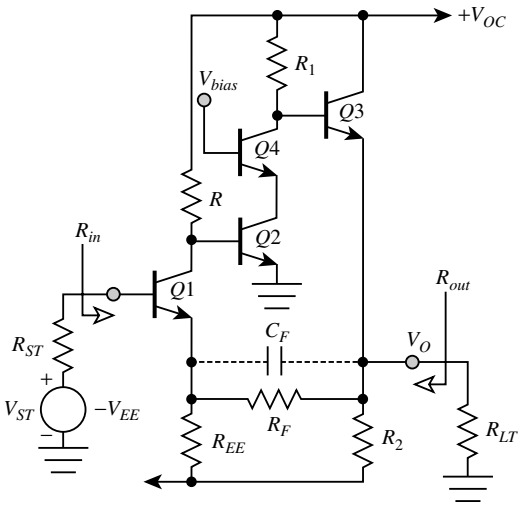


FIGURE 12.6 Buffered series-shunt feedback amplifier with common base cascode compensation of the common emitter amplifier formed by transistor $Q2$. A feedback zero is introduced by the capacitance C_F to achieve acceptable closed-loop damping.

Using the concepts evoked by (12.17) and (12.18), the driving-point input and output impedances can now be determined. In a typical realization of the buffered series-shunt feedback amplifier, the resistance, R_2, in Figure 12.4 is very large because it is manifested as the output resistance of a common base current sink that is employed to stabilize the operating point of transistor $Q3$. For this situation, and assuming the resistance R_1 is large, the resultant driving-point input resistance is larger than its predecessor input resistance by a factor of approximately $(\beta + 1)$. Similarly, it is easy to show that for large R_1 and large R_2, the driving-point output resistance is smaller than that predicted by (12.18) by a factor approaching $(\beta + 1)$.

Although the emitter follower output stage in Figure 12.4 all but eliminates feed-forward signal transmission through the feedback network and increases both the driving point input resistance and output conductance, a potential bandwidth penalty is paid by its incorporation into the basic series-shunt feedback cell. The fundamental problem is that if R_1 is too large, potentially significant Miller multiplication of the base-collector transition capactiance of transistor $Q2$ materializes. The resultant capacitive loading at the collector of transistor $Q1$ is exacerbated by large R, which may produce a dominant pole at a frequency that is too low to satisfy closed-loop bandwidth requirements. The bandwidth problem may be mitigated by coupling resistance R_1 to the collector of $Q2$ through a common base cascode. This stage appears as transistor $Q4$ in Figure 12.6.

Unfortunately, the use of the common base cascode indicated in Figure 12.6 may produce an open-loop amplifier with transfer characteristics that do not emulate a dominant pole response. In other words, the frequency of the compensated pole established by capacitive loading at the collector of transistor $Q1$ may be comparable to the frequencies of poles established elsewhere in the circuit, and particularly at the base node of transistor $Q1$. In this event, frequency compensation aimed toward achieving acceptable closed-loop damping can be implemented by replacing the feedback resistor R_F with the parallel combination of R_F and a feedback capacitance, say C_F, as indicated by the dashed branch in Figure 12.6. The resultant frequency-domain feedback factor $f(s)$ is

$$f(s) = f\left[\frac{1+\frac{s}{z}}{1+\frac{fs}{z}}\right] \tag{12.24}$$

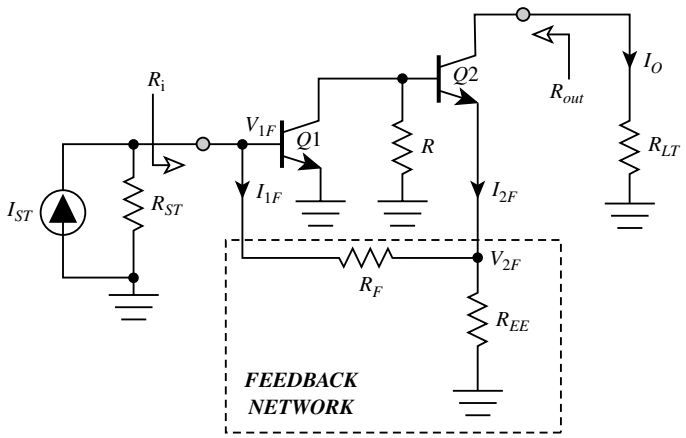
where f is the feedback factor given by (12.7) and z is the frequency of the introduced compensating zero, is

$$z = \frac{1}{R_F C_F} \tag{12.25}$$

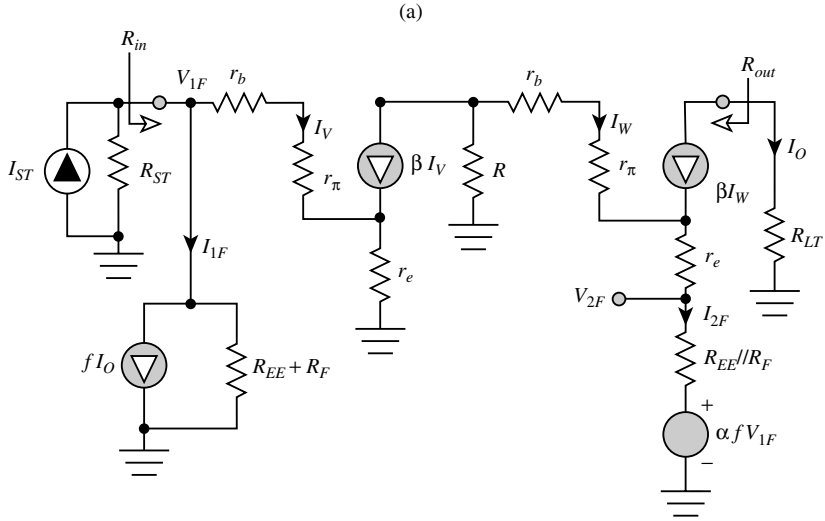
The pole in (12.24) is inconsequential if the closed-loop amplifier bandwidth B_{cl} satisfies the restriction, $f B_{cl} R_F C_F = B_{cl}(R_{EE} \| R_F) C_F \ll 1$.

12.3 Shunt-Series Feedback Amplifier

Although the series-shunt circuit functions as a voltage amplifier, the shunt-series configuration (see the ac schematic diagram depicted in Figure 12.7(a)) is best suited as a current amplifier. In the subject circuit, the $Q2$ emitter current, which is a factor of $(1/\alpha)$ of the output signal current, I_O, is sampled by the feedback network formed of the resistances, R_{EE} and R_F. The sampled current is fed back as a current in shunt with the amplifier input port. Because output current is fed back as a current to a current-



(a)



(b)

FIGURE 12.7 (a) AC schematic diagram of a bipolar shunt-series feedback amplifier. (b) Low-frequency small-signal equivalent circuit of the feedback amplifier.

driven input port, the resultant driving point output resistance is large, and the driving-point input resistance is small. These characteristics allow for a closed-loop current gain, $G_I(R_{ST}, R_{LT}) = I_o/I_{ST}$, that is relatively independent of source and load resistances and insensitive to transistor parameters.

In the series-shunt amplifier, *h* parameters were selected to model the feedback network because the topology of an *h*-parameter equivalent circuit is, similar to the amplifier in which the feedback network is embedded, a series shunt, or Thévenin–Norton, topology, In analogous train of thought compels the use of *g*-parameters to represent the feedback network in Figure 12.7(a). With reference to the branch variables defined in the schematic diagram,

$$\begin{bmatrix} I_{1F} \\ V_{2F} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{R_{EE}+R_F} & -\dfrac{R_{EE}}{R_{EE}+R_F} \\ \dfrac{R_{EE}}{R_{EE}+R_F} & R_{EE}\|R_F \end{bmatrix} \begin{bmatrix} V_{1F} \\ I_{2F} \end{bmatrix} \tag{12.26}$$

Noting that the feedback network current, I_{2F}, relates to the amplifier output current, I_O, in accordance with

$$I_{2F} = -\frac{I_O}{\alpha} \tag{12.27}$$

and letting the feedback factor, *f*, be

$$f = \frac{1}{\alpha}\left(\frac{R_{EE}}{R_{EE}+R_F}\right) \tag{12.28}$$

the small-signal equivalent circuit of shunt-series feedback amplifier becomes the network diagrammed in Figure 12.7(b). Note that the voltage controlled voltage source, $\alpha f V_{1F}$, models the feed-forward transfer mechanism of the feedback network, where the controlling voltage, V_{1F}, is

$$V_{1F} = \left[r_b + r_\pi + (\beta+1)r_c\right]I_V = (\beta+1)r_{ib}I_V \tag{12.29}$$

An analysis of the model in Figure 12.7(b) confirms that the second-stage, signal-base current I_w relates to the first-stage, signal-base current I_v as

$$\frac{I_W}{I_V} = -\frac{\alpha\left(R+fr_{ib}\right)}{r_{ib}+R_{EE}\|R_F+(1-\alpha)R} \tag{12.30}$$

For

$$f \ll \frac{R}{r_{ib}} \tag{12.31}$$

which offsets feed-forward effects,

$$\frac{I_W}{I_V} \approx -\frac{\alpha R}{r_{ib}+R_{EE}\|R_F+(1-\alpha)R} \triangleq -K_r \tag{12.32}$$

Observe that the constant K_r tends toward β for large *R*, as can be verified by an inspection of Figure 12.7(b).

Using (12.32), the open-loop current gain, found by setting f to zero, is

$$G_{\mathrm{IO}}\left(R_{\mathrm{ST}},R_{\mathrm{LT}}\right)=\left.\frac{I_O}{I_{\mathrm{ST}}}\right|_{f=0}=\alpha K_r\left\{\frac{R_{\mathrm{ST}}\left\|\left(R_{\mathrm{EE}}+R_F\right)\right.}{r_{ib}+\left(1-\alpha\right)\left[R_{\mathrm{ST}}\left\|\left(R_{\mathrm{EE}}+R_F\right)\right.\right]}\right\} \tag{12.33}$$

and, recalling (12.28), the loop gain T is

$$\begin{aligned}T=fG_{\mathrm{IO}}\left(R_{\mathrm{ST}},R_{\mathrm{LT}}\right)&=\frac{1}{\alpha}\left(\frac{R_{\mathrm{EE}}}{R_{\mathrm{EE}}+R_F}\right)G_{\mathrm{IO}}\left(R_{\mathrm{ST}},R_{\mathrm{LT}}\right)\\&=K_r\left(\frac{R_{\mathrm{EE}}}{R_{\mathrm{EE}}+R_F+R_{\mathrm{ST}}}\right)\left\{\frac{R_{\mathrm{ST}}}{r_{ib}+\left(1-\alpha\right)\left[R_{\mathrm{ST}}\left\|\left(R_{\mathrm{EE}}+R_F\right)\right.\right]}\right\}\end{aligned} \tag{12.34}$$

By inspection of the model in Figure 12.7(b), the open-loop input resistance, R_{ino}, is

$$R_{\mathrm{ino}}=\left(R_{\mathrm{EE}}+R_F\right)\left\|\left[\left(\beta+1\right)r_{ib}\right]\right. \tag{12.35}$$

and, within the context of an infinitely large Early resistance, the open-loop output resistance, R_{outo}, is infinitely large.

The closed-loop current gain of the shunt-series feedback amplifier is now found to be

$$G_{\mathrm{I}}\left(R_{\mathrm{ST}},R_{\mathrm{LT}}\right)=\frac{G_{\mathrm{IO}}\left(R_{\mathrm{ST}},R_{\mathrm{LT}}\right)}{1+T}\approx\alpha\left(1+\frac{R_F}{R_{\mathrm{EE}}}\right) \tag{12.36}$$

where the indicated approximation exploits the presumption that the loop gain T is much larger than one. As a result of the large loop-gain assumption, note that the closed-loop gain is independent of the source and load resistances and is invulnerable to uncertainties and perturbations in transistor parameters. The closed-loop output resistance, which exceeds its open-loop counterpart, remains infinitely large. Finally, the closed-loop driving point input resistance of the shunt-series amplifier is

$$R_{\mathrm{in}}=\frac{R_{\mathrm{ino}}}{1+fG_{\mathrm{IO}}\left(\infty,R_{\mathrm{LT}}\right)}\approx\left(1+\frac{R_F}{R_{\mathrm{EE}}}\right)\frac{r_{ib}}{K_r} \tag{12.37}$$

12.4 Shunt-Shunt Feedback Amplifier

Circuit Modeling and Analysis

The ac schematic diagram of the third type of single-loop feedback amplifier, the shunt-shunt triple, is drawn in Figure 12.8(a). A cascade interconnection of three transistors $Q1$, $Q2$, and $Q3$, forms the open loop, while the feedback subcircuit is the single resistance, R_F. This resistance samples the output voltage, V_O, as a current fed back to the input port. Output voltage is fed back as a current to a current-driven input port, so both the driving point input and output resistances are very small. Accordingly, the circuit operates best as a transresistance amplifier in that its closed-loop transresistance, $R_M\left(R_{\mathrm{ST}},R_{\mathrm{LT}}\right)=V_O/I_{\mathrm{ST}}$, is nominally invariant with source resistance, load resistance, and transistor parameters.

The shunt-shunt nature of the subject amplifier suggests the propriety of y-parameter modeling of the feedback network. For the electrical variables indicated in Figure 12.8(a),

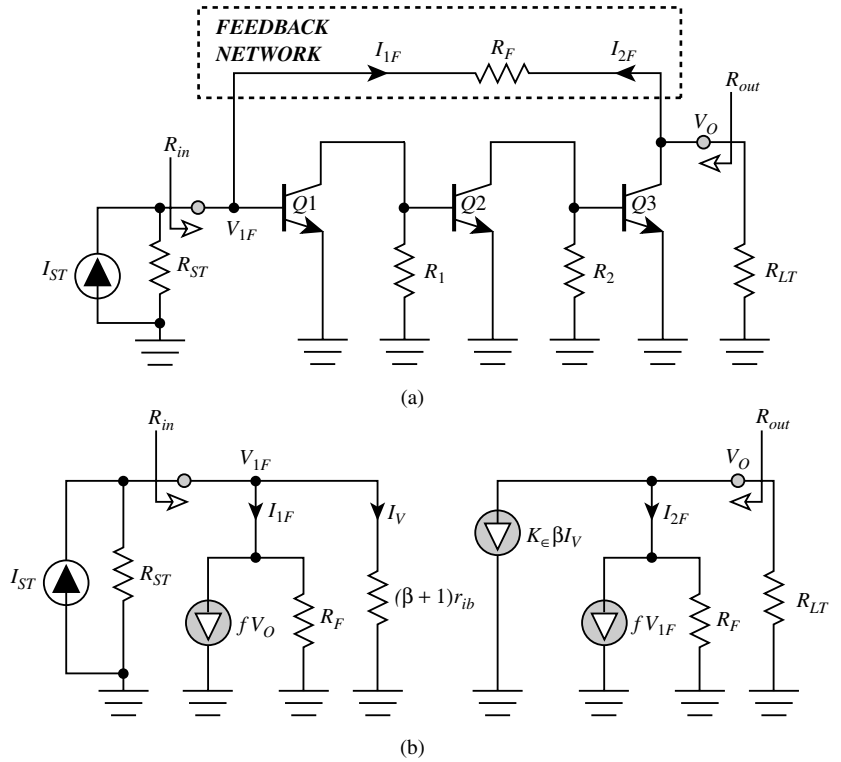


FIGURE 12.8 (a) AC schematic diagram of a bipolar shunt-shunt feedback amplifier. (b) Low-frequency small-signal equivalent circuit of the feedback amplifier.

$$\begin{bmatrix} I_{1F} \\ I_{2F} \end{bmatrix} = \begin{bmatrix} \dfrac{1}{R_F} & -\dfrac{1}{R_F} \\ -\dfrac{1}{R_F} & \dfrac{1}{R_F} \end{bmatrix} \begin{bmatrix} V_{1F} \\ V_O \end{bmatrix} \tag{12.38}$$

which implies that a resistance, R_F, loads both the input and the output ports of the open-loop three-stage cascade. The short-circuit admittance relationship in (12.38) also suggests a feedback factor, f, given by

$$f = \frac{1}{R_F} \tag{12.39}$$

The foregoing observations and the small-signal modeling experience gained with the preceding two feedback amplifiers lead to the equivalent circuit submitted in Figure 12.8(b). For analytical simplicity, the model reflects the assumption that all three transistors in the open loop have identical small-signal parameters. Moreover, the constant, K_ε, which symbolizes the ratio of the signal base current flowing into transistor $Q3$ to the signal base current conducted by transistor $Q1$, is given by

$$K_\in = \left[\frac{\alpha R_1}{r_{ib} + (1-\alpha)R_1}\right]\left[\frac{\alpha R_2}{r_{ib} + (1-\alpha)R_2}\right] \tag{12.40}$$

Finally, the voltage-controlled current source, fV_{1F}, accounts for feed-forward signal transmission through the feedback network. If such feed-forward is to be negligible, the magnitude of this controlled current

must be significantly smaller than $K_\varepsilon \beta I_v$, a current that emulates feed-forward through the open-loop amplifier. Noting that the input port voltage, V_{1F}, in the present case remains the same as that specified by (12.29), negligible feed-forward through the feedback network mandates

$$R_F \gg \frac{r_{ib}}{\alpha K_\varepsilon} \tag{12.41}$$

Because the constant K_ε in (12.40) tends toward β^2 if R_1 and R_2 are large resistances, (12.41) is relatively easy to satisfy.

With feed-forward through the feedback network ignored, an analysis of the model in Figure 12.8(b) provides an open-loop transresistance, $R_{MO}(R_{ST}, R_{LT})$, of

$$R_{MO}\left(R_{ST}, R_{LT}\right) = -\alpha K_\varepsilon \left[\frac{R_F \| R_{ST}}{r_{ib}\left(1-\alpha\right)\left(R_F \| R_{ST}\right)}\right]\left(R_F \| R_{LT}\right) \tag{12.42}$$

while the loop gain is

$$\begin{aligned} T = fR_{MO}\left(R_{ST}, R_{LT}\right) &= -\frac{R_{MO}\left(R_{ST}, R_{LT}\right)}{R_F} \\ &= \alpha K_\varepsilon \left[\frac{R_{ST}}{R_{ST}+R_F}\right]\left[\frac{R_F \| R_{ST}}{r_{ib}\left(1-\alpha\right)\left(R_F \| R_{ST}\right)}\right] \end{aligned} \tag{12.43}$$

For $T \gg 1$, the corresponding closed-loop transresistance $R_M(R_{ST}, R_{LT})$ is

$$R_M\left(R_{ST}, R_{LT}\right) = \frac{R_{MO}\left(R_{ST}, R_{LT}\right)}{1+T} \approx -R_F \tag{12.44}$$

Finally, the approximate driving-point input and output resistances are, respectively,

$$R_{in} \approx \left(\frac{r_{ib}}{\alpha K_\varepsilon}\right)\left(1+\frac{R_F}{R_{LT}}\right) \tag{12.45}$$

$$R_{out} \approx \left[\frac{r_{ib}+\left(1-\alpha\right)\left(R_F \| R_{ST}\right)}{\alpha K_\varepsilon}\right]\left(1+\frac{R_F}{R_{ST}}\right) \tag{12.46}$$

Design Considerations

Because the shunt-shunt triple uses three gain stages in the open-loop amplifier, its loop gain is significantly larger than the loop gains provided by either of the previously considered feedback cells. Accordingly, the feedback triple affords superior desensitization of the closed-loop gain with respect to transistor parameters and source and load resistances; but the presence of a cascade of three common emitter gain stages in the open loop of the amplifier complicates frequency compensation and limits the 3-dB bandwidth. The problem is that, although each common emitter stage approximates a dominant pole amplifier, none of the critical frequencies in the cluster of poles established by the cascade interconnection of these units is likely to be dominant. The uncompensated closed loop is therefore predisposed to unacceptable underdamping, thereby making compensation via an introduced feedback zero difficult.

At least three compensation techniques can be exploited to optimize the performance of the shunt-shunt feedback amplifier [3], [7–9]. The first of these techniques entail pole splitting of the open-loop

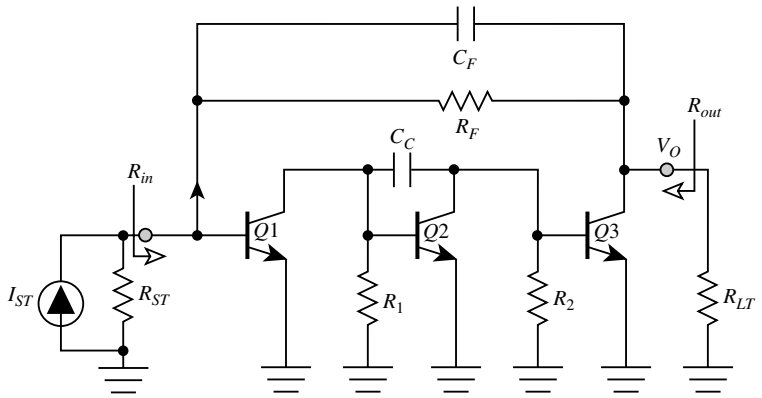


FIGURE 12.9 AC schematic diagram of a frequency compensated shunt-shunt triple. The capacitance, C_c, achieves open-loop pole splitting, while the capacitance, C_F, implements a compensating feedback network zero.

interstage through the introduction of a capacitance, C_c, between the base and the collector terminals of transistor $Q2$, as depicted in the ac schematic diagram of Figure 12.9. In principle, pole splitting can be invoked on any one of the three stages of the open loop; but pole splitting of the interstage is most desirable because such compensation of the first stage proves effective only for large source resistance. Moreover, the resultant dominant pole becomes dependent on the source termination. On the other hand, pole splitting of the third stage produces a dominant pole that is sensitive to load termination. In conjunction with pole splitting, a feedback zero can be introduced, if necessary, to increase closed-loop damping by replacing the feedback resistance, R_F, by the parallel combination of R_F and a feedback capacitance, C_F, as illustrated in Figure 12.9. This compensation produces left-half-plane zero in the feedback factor at $s = -(1/R_F)$.

A second compensation method broadbands the interstage of the open-loop amplifier through local current feedback introduced by the resistance, R_X, in Figure 12.10. Simultaneously, the third stage is broadbanded by way of a common base cascode transistor $Q4$. Because emitter degeneration of the interstage reduces the open-loop gain, an emitter follower (transistor $Q5$) is embedded between the feedback network

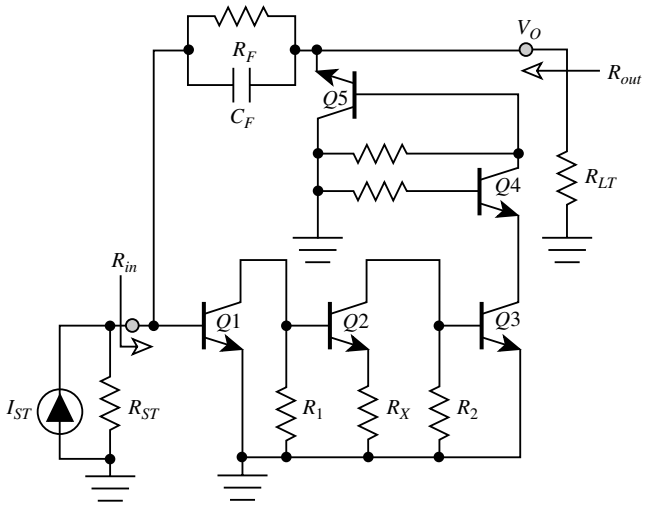


FIGURE 12.10 AC schematic diagram of an alternative compensation scheme for the shunt-shunt triple. Transistor $Q2$ is broadbanded by the emitter degeneration resistance R_X and transistor $Q3$ is broadbanded by the common base cascode transistor $Q4$. The emitter follower transistor, $Q5$, minimizes feed-forward signal transmission through the feedback network.

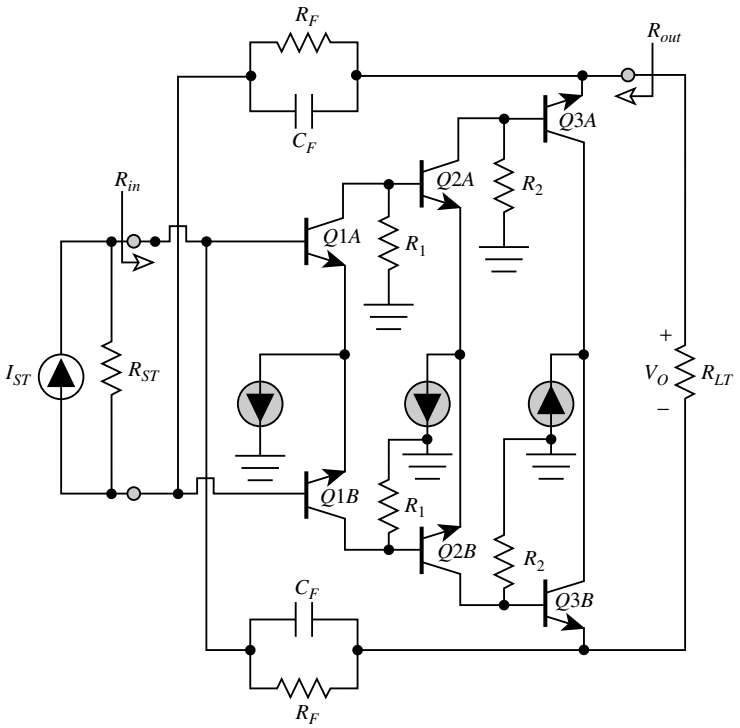


FIGURE 12.11 AC schematic diagram of a differential realization of the compensated shunt-shunt feedback amplifier. The balanced stage boasts improved bandwidth over its single-ended counterpart because of its use of only two high-gain stages in the open loop. The emitter follower pair *Q*3A and *Q*3B diminishes feed-forward transmission through the feedback network composed of the shunt interconnection of resistor R_F with capacitor C_F.

and the output port of the open-loop third stage. As in the case of the series-shunt feedback amplifier, the first-order effect of this emitter follower is to increase feed-forward signal transmission through the open-loop amplifier by a factor that approaches $(\beta + 1)$.

A final compensation method is available if shunt-shunt feedback is implemented as the balanced differential architecture (see the ac schematic diagram offered in Figure 12.11). By exploiting the antiphase nature of opposite collectors in a balanced common emitter topology, a shunt-shunt feedback amplifier can be realized with only two gain stages in the open loop. The resultant closed loop 3-dB bandwidth is invariably larger than that of its three-stage single-ended counterpart, because the open loop is now characterized by only two, as opposed to three, fundamental critical frequencies. Because the forward gain implicit to two amplifier stages is smaller than the gain afforded by three stages of amplification, a balanced emitter follower (transistors *Q*3*A* and *Q*3*B*) is incorporated to circumvent the deleterious relative effects of feed-forward signal transmission through the feedback network.

12.5 Series-Series Feedback Amplifier

Figure 12.12(a) is the ac schematic diagram of the series-series feedback amplifier. Three transistors, *Q*1, *Q*2, and *Q*3, are embedded in the open-loop amplifier, while the feedback subcircuit is the wye configuration formed of the resistances R_X, R_Y, and R_Z. Although it is possible to realize series-series feedback via emitter degeneration of a single-stage amplifier, the series-series triple offers substantially more loop gain and thus better desensitization of the forward gain with respect to both transistor parameters and source and load terminations.

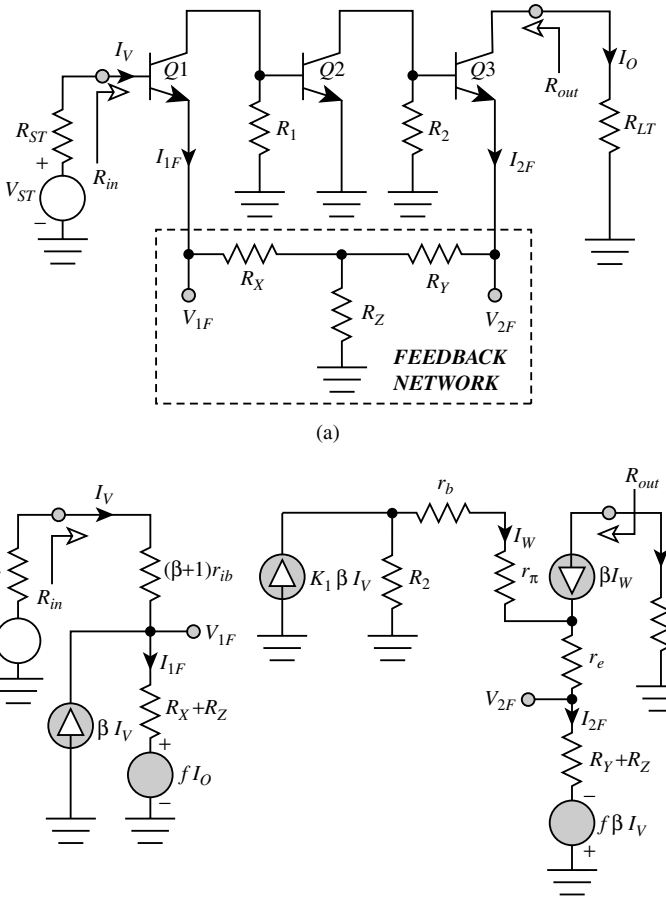


FIGURE 12.12 (a) AC schematic diagram of a bipolar series-series feedback amplifier. (b) Low-frequency, small-signal equivalent circuit of the feedback amplifier.

In Figure 12.12(a), the feedback wye senses the $Q3$ emitter current, which is a factor of $(1/\alpha)$ of the output signal current I_o. This sampled current is fed back as a voltage in series with the emitter of $Q1$. Because output current is fed back as a voltage to a voltage-driven input port, both the driving point input and output resistances are large. The circuit is therefore best suited as a transconductance amplifier in the sense that for large loop gain, its closed-loop transconductance, $G_M(R_{ST}, R_{LT}) = I_O/V_{ST}$, is almost independent of the source and load resistances.

The series-series topology of the subject amplifier conduces z-parameter modeling of the feedback network. Noting the electrical variables delineated in the diagram of Figure 12.12(a),

$$\begin{bmatrix} V_{1F} \\ V_{2F} \end{bmatrix} = \begin{bmatrix} R_X + R_Z & R_Z \\ R_Z & R_Y + R_Z \end{bmatrix} \begin{bmatrix} I_{1F} \\ I_{2F} \end{bmatrix} \tag{12.47}$$

Equation (12.47) suggests that the open-circuit feedback network resistances loading the emitters of transistors $Q1$ and $Q3$ are $(R_X + R_Z)$ and $(R_Y + R_Z)$, respectively, and the voltage fed back to the emitter of transistor $Q1$ is $R_Z I_{2F}$. Because the indicated feedback network current I_{2F} is $(-I_O/\alpha)$, this fed back voltage is equivalent to $(-R_Z I_O/\alpha)$, which suggests a feedback factor, f, of

$$f = \frac{R_Z}{\alpha} \tag{12.48}$$

Finally, the feed-forward through the feedback network if $R_Z I_{1F}$. Because I_{1F} relates to the signal base current I_V flowing into transistor $Q1$ by $I_{1F} = (\beta + 1)I_v$, this feed-forward voltage is also expressible as $(-f\,\beta I_v)$. The foregoing observations and the hybrid-pi method of a bipolar junction transistor produce the small-signal model depicted in Figure 12.12(b). In this model, all transistors are presumed to have identical corresponding small-signal parameters, and the constant, K_1, is

$$K_1 = \frac{\alpha R_1}{r_{ib} + (1-\alpha)R_1} \tag{12.49}$$

An analysis of the model of Figure 12.12(b) confirms that the ratio of the signal current, I_W, flowing into the base of transistor $Q3$ to the signal base current, I_V, of transistor $Q1$ is

$$\frac{I_W}{I_V} = \frac{\alpha K_1 R_2 \left(1 + \dfrac{f}{K_1 R_2}\right)}{r_{ib} + R_Y + R_Z + (1-\alpha)R_2} \tag{12.50}$$

This result suggests that feed-forward effects through the feedback network are negligible if $|f| \ll K_1 R_2$, which requires

$$R_Z \ll \alpha K_1 R_2 \tag{12.51}$$

In view of the fact that the constant, K_1, approaches β for large values of the resistance, R_1, (12.51) is not a troublesome inequality. Introducing a second constant, K_2, such that

$$K_2 \triangleq \frac{\alpha R_2}{r_{ib} + R_Y + R_Z + (1-\alpha)R_2} \tag{12.52}$$

the ratio I_W/I_V in (12.50) becomes

$$\frac{I_W}{I_V} \approx K_1 K_2 \tag{12.53}$$

assuming (12.51) is satisfied.

Given the propriety of (12.50) and using (12.53) the open-loop transconductance, $G_{MO}(R_{ST}, R_{LT})$ is found to be

$$G_{MO}(R_{ST}, R_{LT}) = -\left\{\frac{\alpha K_1 K_2}{r_{ib} + R_X + R_Z + (1-\alpha)R_{ST}}\right\} \tag{12.54}$$

and recalling (12.48), the loop gain T is

$$T = -\left(\frac{R_Z}{\alpha}\right) G_{MO}(R_{ST}, R_{LT}) = \frac{K_1 K_2 R_Z}{r_{ib} + R_X + R_Z + (1-\alpha)R_{ST}} \tag{12.55}$$

It follows that for $T \gg 1$, the closed-loop transconductance is

$$G_M(R_{ST}, R_{LT}) = \frac{G_{MO}(R_{ST}, R_{LT})}{1+T} \approx -\frac{\alpha}{R_Z} \tag{12.56}$$

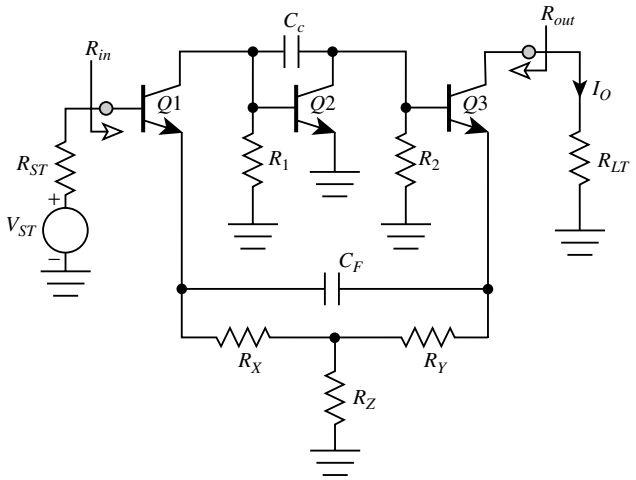


FIGURE 12.13 AC schematic diagram of a frequency compensated series-series feedback triple. The capacitance, C_c, achieves pole splitting in the open-loop configuration, while the capacitance, C_F, introduces a zero in the feedback factor of the closed-loop amplifier.

The Early resistance is large enough to justify its neglect, so the open-loop, and thus the closed-loop, driving-point output resistances are infinitely large. On the other hand, the closed-loop driving point input resistance R_{in} can be shown to be

$$R_{\text{in}} = R_{\text{ino}}\left[1 + f\,G_{\text{MO}}\left(0, R_{\text{LT}}\right)\right] \approx \left(\beta + 1\right)K_1 K_2 R_Z \tag{12.57}$$

Similar to its shunt-shunt counterpart, the series-series feedback amplifier uses three open-loop gain stages to produce large loop gain. However, also similar to the shunt-shunt triple, frequency compensation via an introduced feedback zero is difficult unless design care is exercised to realize a dominant pole open-loop response. To this end, the most commonly used compensation is pole splitting in the open loop, combined, if required, with the introduction of a zero in the feedback factor. The relevant ac schematic diagram appears in Figure 12.13 where the indicated capacitance, C_c, inserted across the base-collector terminals of transistor $Q3$ achieves the aforementioned pole splitting compensation. The capacitance, C_F, in Figure 12.13 delivers a frequency-dependent feedback factor, $f\,(s)$ of

$$f(s) = f\left[\frac{1 + \dfrac{s}{z}}{1 + \dfrac{s}{z}\left(\dfrac{R_Z}{R_Z + R_X \| R_Y}\right)}\right] \tag{12.58}$$

where the frequency z of the introduced zero derives from

$$\frac{1}{z} = \left(R_X + R_Y\right)\left(1 + \frac{R_X \| R_Y}{R_Z}\right)C_F \tag{12.59}$$

The corresponding pole in (12.58) is insignificant if the closed-loop amplifier is designed for a bandwidth, B_{cl} that satisfies the inequality, $B_{\text{cl}}(R_X + R_Y)C_F \ll 1$.

As is the case with shunt-shunt feedback, an alternative frequency compensation scheme is available if series-series feedback is implemented as a balanced differential architecture. The pertinent ac schematic

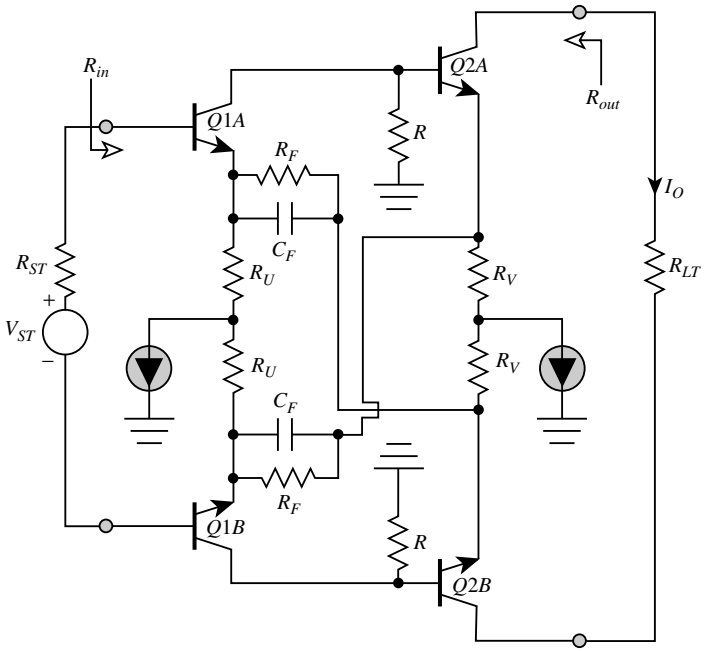


FIGURE 12.14 AC schematic diagram of a balanced differential version of the series-series feedback amplifier. The circuit utilizes only two, as opposed to three, gain stages in the open loop.

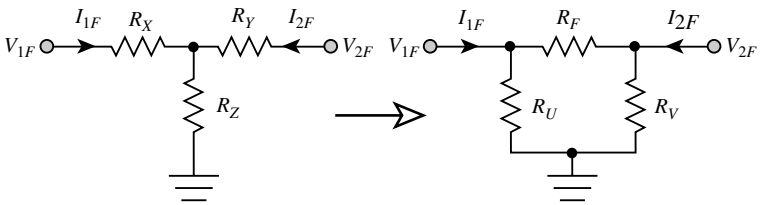


FIGURE 12.15 Transformation of the wye feedback subcircuit used in the amplifier of Figure 12.13 to the delta subcircuit exploited in Figure 12.14. The resistance transformation equations are given by (12.60)–(12.62).

diagram, inclusive of feedback compensation, appears in Figure 12.14. This diagram exploits the fact that the feedback wye consisting of the resistances, R_X, R_Y, and R_Z as utilized in the single-ended configurations of Figures 12.12(a) and 12.13 can be transformed into the feedback delta of Figure 12.15. The terminal volt-ampere characteristics of the two networks in Figure 12.15 are identical, provided that the delta subcircuit elements, R_F, R_U, and R_V, are chosen in accordance with

$$R_F = \left(R_X + R_Y\right)\left(1 + \frac{R_X \| R_Y}{R_Z}\right) \tag{12.60}$$

$$\frac{R_U}{R_F} = \frac{R_Z}{R_Y} \tag{12.61}$$

$$\frac{R_V}{R_F} = \frac{R_Z}{R_X} \tag{12.62}$$

12.6 Dual-Loop Feedback

As mentioned previously, a simultaneous control of the driving point I/O resistances, as well as the closed-loop gain, mandates the use of dual global loops comprised of series and shunt feedback signal paths. The two global dual-loop feedback architectures are the **series-series/shunt-shunt feedback amplifier** and the **series-shunt/shunt-series feedback amplifier**. In the following subsections, both of these units are studied by judiciously applying the relevant analytical results established earlier for pertinent single-loop feedback architectures. The ac schematic diagrams of these respective circuit realizations are provided, and engineering design considerations are offered.

Series-Series/Shunt-Shunt Feedback Amplifier

Figure 12.16 is a behavioral abstraction of the series-series/shunt-shunt feedback amplifier. Two port z parameters are used to model the series-series feedback subcircuit, for which feed-forward is tacitly ignored and the feedback factor associated with its current controlled voltage source is f_{ss}. On the other hand, y parameters model the shunt-shunt feedback network, where the feedback factor relative to its voltage controlled current source is f_{pp}. As in the series-series network, feed-forward in the shunt-shunt subcircuit is presumed negligible. The four-terminal amplifier around which the two feedback units are connected has an open loop (meaning $f_{ss} = 0$ and $f_{pp} = 0$, but with the loading effects of both feedback circuits considered) transconductance of $G_{\mathrm{MO}}(R_{\mathrm{ST}}, R_{\mathrm{LT}})$.

With f_{pp} set to zero to deactivate shunt-shunt feedback, the resultant series-series feedback network is a transconductance amplifier with a closed-loop transconductance, $G_{\mathrm{MS}}(R_{\mathrm{ST}}, R_{\mathrm{LT}})$, is

$$G_{\mathrm{MS}}\left(R_{\mathrm{ST}}, R_{\mathrm{LT}}\right) = \frac{I_O}{V_{\mathrm{ST}}} = \frac{G_{\mathrm{MO}}(R_{\mathrm{ST}}, R_{\mathrm{LT}})}{1 + f_{ss} G_{\mathrm{MO}}\left(R_{\mathrm{ST}}, R_{\mathrm{LT}}\right)} \approx \frac{1}{f_{ss}} \tag{12.63}$$

where the loop gain, $f_{ss}G_{\mathrm{MO}}(R_{\mathrm{ST}}, R_{\mathrm{LT}})$, is presumed much larger than one, and the loading effects of both the series-series feedback subcircuit and the deactivated shunt-shunt feedback network are incorporated

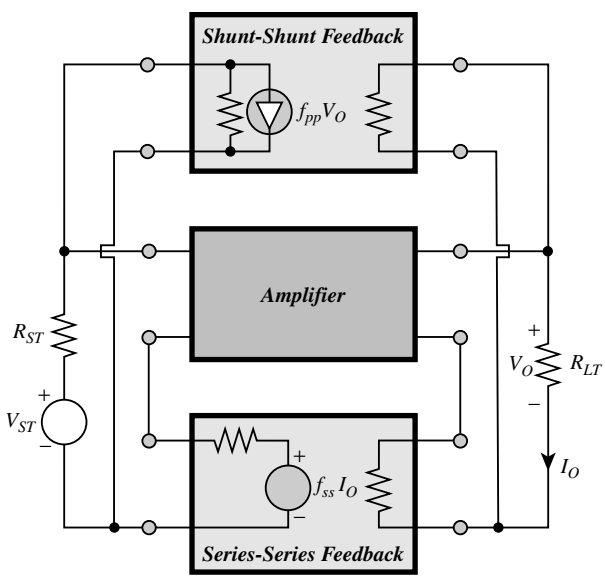


FIGURE 12.16 System-level diagram of a series-series/shunt-shunt dual-loop feedback amplifier. Note that feed-forward signal transmission through either feedback network is ignored.

into $G_{MO}(R_{ST}, R_{LT})$. The transresistance, $R_{MS}(R_{ST}, R_{LT})$, implied by (12.63), which expedites the study of the shunt-shunt component of the feedback configuration, is

$$R_{MS}\left(R_{ST}, R_{LT}\right) = \frac{V_O}{I_{ST}} = R_{ST}R_{LT}\frac{I_O}{V_{ST}} \approx \frac{R_{ST}R_{LT}}{f_{ss}} \tag{12.64}$$

The series-series feedback input and output resistances R_{ins} and R_{outs}, respectively, are large and given by

$$R_{ins} = R_{ino}\left[1 + f_{ss}G_{MO}\left(0, R_{LT}\right)\right] \tag{12.65}$$

and

$$R_{outs} = R_{outo}\left[1 + f_{ss}G_{MO}\left(R_{ST}, 0\right)\right] \tag{12.66}$$

where the zero feedback ($f_{ss} = 0$ and $f_{pp} = 0$) values, R_{ino} and R_{outo}, of these driving point quantities are computed with due consideration given to the loading effects imposed on the amplifier by both feedback subcircuits.

When shunt-shunt feedback is applied around the series-series feedback cell, the configuration becomes a transresistance amplifier. The effective open-loop transresistance is $R_{MS}(R_{ST}, R_{LT})$, as defined by (12.64). Noting a feedback of f_{pp}, the corresponding closed-loop transresistance is

$$R_M\left(R_{ST}, R_{LT}\right) \approx \frac{\dfrac{R_{ST}R_{LT}}{f_{ss}}}{1 + f_{pp}\left(\dfrac{R_{ST}R_{LT}}{f_{ss}}\right)} \tag{12.67}$$

which is independent of amplifier model parameters, despite the unlikely condition of an effective loop gain $f_{pp}R_{ST}R_{LT}/f_{ss}$ that is much larger than one. It should be interjected, however, that (12.67) presumes negligible feed-forward through the shunt-shunt feedback network. This presumption may be inappropriate owing to the relatively low closed-loop gain afforded by the series-series feedback subcircuit. Ignoring this potential problem temporarily, (12.67) suggests a closed-loop voltage gain $A_V\,(R_{ST}, R_{LT})$ of

$$A_V\left(R_{ST}, R_{LT}\right) = \frac{V_O}{V_S} = \frac{R_M\left(R_{ST}, R_{LT}\right)}{R_{ST}} \approx \frac{R_{LT}}{f_{ss} + f_{pp}R_{ST}R_{LT}} \tag{12.68}$$

The closed-loop, driving-point output resistance R_{out}, can be straightforwardly calculated by noting that the open circuit ($R_{LT} \to \infty$) voltage gain, A_{VO}, predicted by (12.68) is $A_{VO} = 1/f_{pp}\,R_{ST}$. Accordingly, (12.68) is alternatively expressible as

$$A_V\left(R_{ST}, R_{LT}\right) \approx A_{VO}\left(\frac{R_{LT}}{R_{LT} + \dfrac{f_{ss}}{f_{pp}R_{ST}}}\right) \tag{12.69}$$

Because (12.69) is a voltage divider relationship stemming from a Thévenin model of the output port of the dual-loop feedback amplifier, as delineated in Figure 12.17, it follows that the driving-point output resistance is

$$R_{out} \approx \frac{f_{ss}}{f_{pp}R_{ST}} \tag{12.70}$$

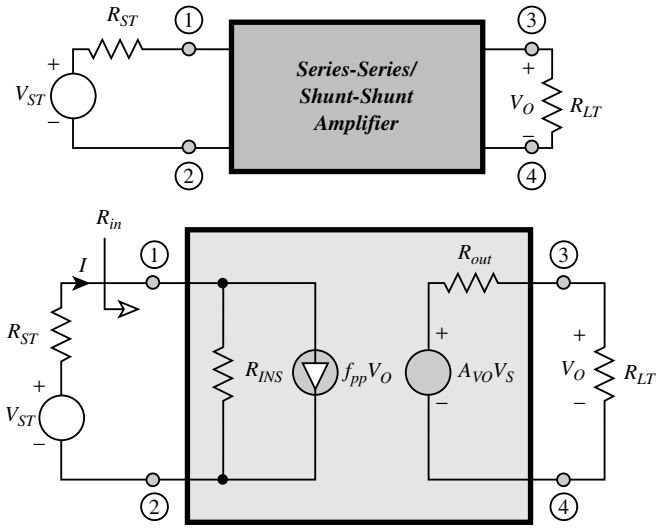


FIGURE 12.17 Norton equivalent input and Thévenin equivalent output circuits for the series-series/shunt-shunt dual-loop feedback amplifier.

Observe that, similar to the forward gain characteristics, the driving-point output resistance is nominally insensitive to changes and other uncertainties in open-loop amplifier parameters. Moreover, this output resistance is directly proportional to the ratio f_{ss}/f_{pp} of feedback factors. As is illustrated in preceding sections, the individual feedback factors, and thus the ratio of feedback factors, is likely to be proportional to a ratio of resistances. In view of the fact that resistance ratios can be tightly controlled in a monolithic fabrication process, R_{out} in (12.70) is accurately prescribed for a given source termination.

The driving-point input resistance R_{in} can be determined from a consideration of the input port component of the system level equivalent circuit depicted in Figure 12.17. This resistance is the ratio of V_{ST} to I, under the condition of $R_S = 0$. With $R_S = 0$, (12.68) yields $V_O = R_{\text{LT}}V_{\text{ST}}/f_{ss}$ and thus, Kirchhoff's voltage law (KVL) applied around the input port of the model at hand yields

$$R_{\text{in}} = \frac{R_{\text{ins}}}{1 + \dfrac{f_{pp}R_{\text{LT}}R_{\text{ins}}}{f_{ss}}} \approx \frac{f_{ss}}{f_{pp}R_{\text{LT}}} \tag{12.71}$$

where the "open-loop" input resistance R_{ins}, defined by (12.65), is presumed large. Similar to the driving-point output resistance of the series-series/shunt-shunt feedback amplifier, the driving-point input resistance is nominally independent of open-loop amplifier parameters.

It is interesting to observe that the input resistance in (12.71) is inversely proportional to the load resistance by the same factor (f_{ss}/f_{pp}) that the driving-point output resistance in (12.70) is inversely proportional to the source resistance. As a result,

$$\frac{f_{ss}}{f_{pp}} \approx R_{\text{in}}R_{\text{LT}} \equiv R_{\text{out}}R_{\text{ST}} \tag{12.72}$$

Thus, in addition to being stable performance indices for well-defined source and load terminations, the driving-point input and output resistances track one another, despite manufacturing uncertainties and changes in operating temperature that might perturb the individual values of the two feedback factors f_{ss} and f_{pp}.

The circuit property stipulated by (12.72) has immediate utility in the design of wideband communication transceivers and other high-speed signal-processing systems [10–14]. In these and related applications, a cascade of several stages is generally required to satisfy frequency response, distortion, and noise specifications. A convenient way of implementing a cascade interconnection is to force each member of the cascade to operate under the match terminated case of $R_{ST} = R_{in} = R_{LT} = R_{out} \triangleq R$. From (12.72) match terminated operation demands feedback factors selected so that

$$R = \sqrt{\frac{f_{ss}}{f_{pp}}} \tag{12.73}$$

which forces a match terminated closed-loop voltage gain A_V^* of

$$A_V^* \approx \frac{1}{2f_{pp}R} = \frac{1}{2\sqrt{f_{pp}f_{ss}}} \tag{12.74}$$

The ac schematic diagram of a practical, single-ended series-series/shunt-shunt amplifier is submitted in Figure 12.18. An inspection of this diagram reveals a topology that coalesces the series-series and shunt-shunt triples studied earlier. In particular, the wye network formed of the three resistances, R_{Xx}, R_Y, and R_Z, comprises the series-series component of the dual-loop feedback amplifier. The capacitor, C_c, narrowbands the open-loop amplifier to facilitate frequency compensation of the series-series loop through the capacitance, C_{F1}. Compensated shunt feedback of the network is achieved by the parallel combination of the resistance, R_F and the capacitance, C_{F2}. If C_{F1} and C_c combine to deliver a dominant pole series-series feedback amplifier, C_{F2} is not necessary. Conversely, C_{F1} is superfluous if C_{F2} and C_c interact to provide a dominant pole shunt-shunt feedback amplifier. As in the single ended series-series configuration, transistor $Q3$ can be broadbanded via a common base cascode. Moreover, if feedback through the feedback networks poses a problem, an emitter follower can be inserted at the port to which the shunt feedback path and the load termination are incident.

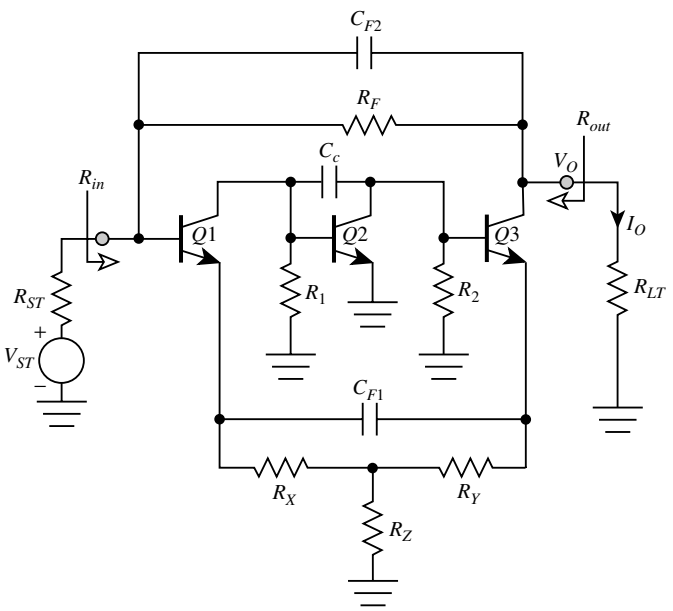


FIGURE 12.18 AC schematic diagram of a frequency-compensated, series-series/shunt-shunt, dual-loop feedback amplifier. The compensation is affected by the capacitances C_{F1} and C_{F2}, while C_c achieves pole splitting in the open-loop amplifier.

A low-frequency analysis of the circuit in Figure 12.18 is expedited by assuming high beta transistors having identical corresponding small-signal model parameters. This analysis, which in contrast to the simplified behavioral analysis, does not ignore the electrical effects of the aforementioned feed-forward through the shunt-shunt feedback network, yields a voltage gain $A_V\,(R_{ST}, R_{LT})$, of

$$A_V\left(R_{ST}, R_{LT}\right) \approx -\left(\frac{R_{in}}{R_{in}+R_{ST}}\right)\left(\frac{R_{LT}}{R_{LT}+R_F}\right)\left(\frac{\alpha R_F}{R_Z}-1\right) \tag{12.75}$$

where the driving-point input resistance of the amplifier R_{in} is

$$R_{in} \approx \frac{R_F + R_{LT}}{1+\dfrac{\alpha R_{LT}}{R_Z}} \tag{12.76}$$

The driving-point output resistance R_{out} is

$$R_{out} \approx \frac{R_F + R_{ST}}{1+\dfrac{\alpha R_{ST}}{R_z}} \tag{12.77}$$

As predicted by the behavioral analysis R_{in}, R_{out}, and $A_V\,(R_{ST}, R_{LT})$, are nominally independent of transistor parameters. Observe that the functional dependence of R_{in} on the load resistance, R_{LT}, is identical to the manner in which R_{out} is related to the source resistance R_{ST}. In particular, $R_{in} \equiv R_{out}$ if $R_{ST} \equiv R_{LT}$. For the match terminated case in which $R_{ST} = R_{in} = R_{LT} = R_{out} \triangleq R$,

$$R \approx \sqrt{\frac{R_F R_Z}{\alpha}} \tag{12.78}$$

The corresponding match terminated voltage gain in (12.75) collapses to

$$A_V^* \approx -\left(\frac{R_F - R}{2R}\right) \tag{12.79}$$

Similar to the series-series and shunt-shunt triples, many of the frequency compensation problems implicit to the presence of three open-loop stages can be circumvented by realizing the series-series/shunt-shunt amplifier as a two-stage differential configuration. Figure 12.19 is the acschematic diagram of a compensated differential series-series/shunt-shunt feedback dual.

Series-Shunt/Shunt-Series Feedback Amplifier

The only other type of global dual loop architecture is the series-shunt/shunt-series feedback amplifier; the behavioral diagram appears in Figure 12.20. The series-shunt component of this system, which is modeled by *h*-parameters, has a negligibly small feed-forward factor and a feedback factor of f_{sp}. Hybrid *g*-parameters model the shunt-series feedback structure, which has a feedback factor of f_{ps} and a presumably negligible feed-forward factor. The four-terminal amplifier around which the two feedback units are connected has an open-loop (meaning $f_{sp} = 0$ and $f_{ps} = 0$, but with the loading effects of both feedback circuits considered) voltage gain of $A_{VO}(R_{ST}, R_{LT})$.

For $f_{ps} = 0$, the series-shunt feedback circuit voltage gain $A_{VS}(R_{ST}, R_{LT})$, is

$$A_{VS}\left(R_{ST}, R_{LT}\right) = \frac{V_O}{V_{ST}} = \frac{A_{VO}\left(R_{ST}, R_{LT}\right)}{1+f_{sp}A_{VO}\left(R_{ST}, R_{LT}\right)} \approx \frac{1}{f_{sp}} \tag{12.80}$$

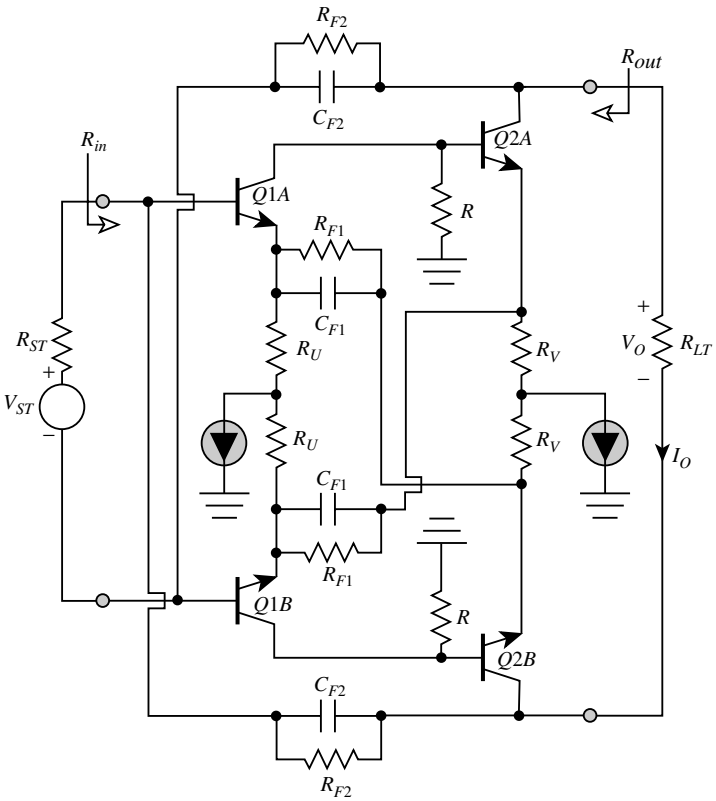


FIGURE 12.19 AC schematic diagram of the differential realization of a compensated series-series/shunt-shunt feedback amplifier.

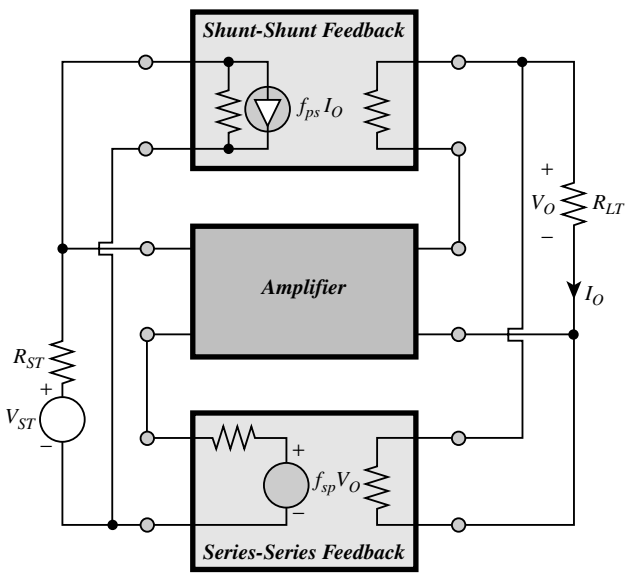


FIGURE 12.20 System level diagram of a series-shunt/shunt-series, dual-loop feedback amplifier. Note that feedforward signal transmission through either feedback network is ignored.

where the approximation reflects an assumption of a large loop gain. When the shunt-series component of the feedback amplifier is activated, the dual-loop configuration functions as a current amplifier. Its effective open-loop transfer function is the current gain, $A_{IS}(R_{ST}, R_{LT})$, established by the series-shunt amplifier; namely,

$$A_{IS}\left(R_{ST}, R_{LT}\right) = \frac{I_O}{I_{ST}} = \left(\frac{R_{ST}}{R_{LT}}\right)\frac{V_O}{V_{ST}} \approx \frac{R_{ST}}{f_{sp}R_{LT}} \tag{12.81}$$

It follows that the current gain, $A_I(R_{ST}, R_{LT})$, of the closed loop is

$$A_I\left(R_{ST}, R_{LT}\right) \approx \frac{\dfrac{R_{ST}}{f_{sp}R_{LT}}}{1 + f_{ps}\left(\dfrac{R_{ST}}{f_{sp}R_{LT}}\right)} = \frac{R_{ST}}{f_{sp}R_{LT} + f_{ps}R_{ST}} \tag{12.82}$$

while the corresponding voltage gain, $A_V(R_{ST}, R_{LT})$, assuming negligible feed-forward through the shunt-series feedback network, is

$$A_V\left(R_{ST}, R_{LT}\right) = \frac{R_{LT}}{R_{ST}} A_I\left(R_{ST}, R_{LT}\right) \approx \frac{R_{LT}}{f_{sp}R_{LT} + f_{ps}R_{ST}} \tag{12.83}$$

Repeating the analytical strategy employed to determine the input and output resistances of the series-series/shunt-shunt configuration, (12.83) delivers a driving-point input resistance of

$$R_{in} \approx \frac{f_{sp}R_{LT}}{f_{ps}} \tag{12.84}$$

and a driving-point output resistance of

$$R_{out} \approx \frac{f_{ps}R_{ST}}{f_{sp}} \tag{12.85}$$

Similar to the forward voltage gain, the driving-point input and output resistances of the series-shunt/shunt-series feedback amplifier are nominally independent of active element parameters. Note, however, that the input resistance is directly proportional to the load resistance by a factor (f_{sp}/f_{ps}), which is the inverse of the proportionality constant that links the output resistance to the source resistance. Specifically,

$$\frac{f_{sp}}{f_{ps}} = \frac{R_{in}}{R_{LT}} = \frac{R_{ST}}{R_{out}} \tag{12.86}$$

Thus, although R_{in} and R_{out} are reliably determined for well-defined load and source terminations, they do not track one another as well as they do in the series-series/shunt-shunt amplifier. Using (12.86), the voltage gain in (12.83) is expressible as

$$A_V\left(R_{ST}, R_{LT}\right) \approx \frac{1}{f_{sp}\left(1 + \sqrt{\dfrac{R_{out}R_{ST}}{R_{in}R_{LT}}}\right)} \tag{12.87}$$

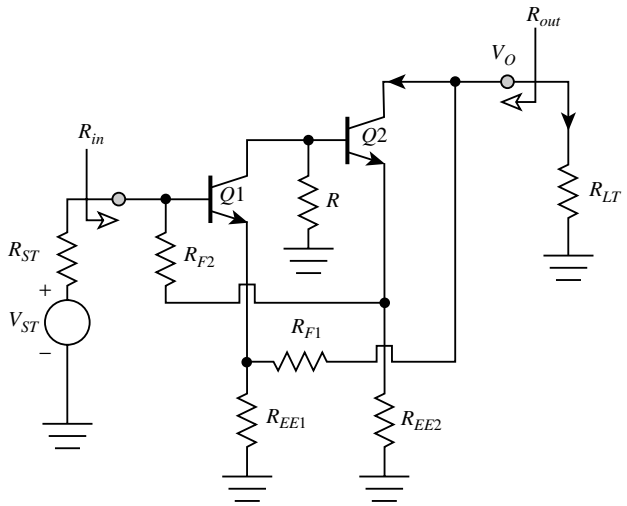


FIGURE 12.21 AC schematic diagram of a series-shunt/shunt-series, dual-loop feedback amplifier.

The simplified ac schematic diagram of a practical series-shunt/shunt-series feedback amplifier appears in Figure 12.21. In this circuit, series-shunt feedback derives from the resistances, R_{EE1} and R_{F1}, and shunt-series feedback is determined by the resistances, R_{EE2} and R_{F2}. Because this circuit topology merges the series-shunt and shunt-series pairs, requisite frequency compensation, which is not shown in the subject figure, mirrors the relevant compensation schemes studied earlier. Note, however, that a cascade of only two open-loop gain stages renders compensation easier to implement and larger 3-dB bandwidths easier to achieve in the series-series/shunt-shunt circuit, which requires three open-loop gain stages for a single-ended application.

For high beta transistors having identical corresponding small-signal model parameters, a low-frequency analysis of the circuit in Figure 12.21 gives a voltage gain of

$$A_V\left(R_{\text{ST}}, R_{\text{LT}}\right) \approx \left(\frac{\alpha R_{\text{in}}}{R_{\text{in}} + \alpha R_S}\right)\left(1 + \frac{R_{\text{F1}}}{R_{\text{EE1}}}\right) \tag{12.88}$$

where the driving-point input resistance, R_{in}, of the subject amplifier is

$$R_{\text{in}} \approx \alpha R_{\text{LT}}\left(\frac{1 + \dfrac{R_{\text{F2}}}{R_{\text{EE2}}}}{1 + \dfrac{R_{\text{F1}}}{R_{\text{EE1}}} + \dfrac{R_{\text{LT}}}{R_{\text{EE1}} \| R_{\text{EE2}}}}\right) \tag{12.89}$$

The driving-point output resistance, R_{out}, is

$$R_{\text{out}} \approx R_{\text{ST}}\left(\frac{1 + \dfrac{R_{\text{F1}}}{R_{\text{EE1}}}}{1 + \dfrac{R_{\text{F2}}}{R_{\text{EE2}}} + \dfrac{R_{\text{ST}}}{R_{\text{EE1}} \| R_{\text{EE2}}}}\right) \tag{12.90}$$

12.7 Summary

This section documents small-signal performance equations, general operating characteristics, and engineering design guidelines for the six most commonly used global feedback circuits. These observations derive from analyses based on the judicious application of signal flow theory to the small-signal model that results when the subject feedback network is supplanted by an appropriate two-port parameter equivalent circuit.

Four of the six fundamental feedback circuits are single-loop architectures.

1. The series-shunt feedback amplifier functions best as a voltage amplifier in that its input resistance is large, and its output resistance is small. Because only two gain stages are required in the open loop, the amplifier is relatively easy to compensate for acceptable closed-loop damping and features potentially large 3-dB bandwidth. A computationally efficient analysis aimed toward determining loop gain, closed-loop gain, I/O resistances, and the condition that renders feed-forward through the feedback network inconsequential is predicated on replacing the feedback subcircuit with its *h*-parameter model.
2. The shunt-series feedback amplifier is a current amplifier in that its input resistance is small, and its output resistance is large. Similar to its series-shunt dual, only two gain stages are required in the open loop. Computationally efficient analyses are conducted by replacing the feedback subcircuit with its *g*-parameter model.
3. The shunt-shunt feedback amplifier is a transresistance signal processor in that both its input and output resistances are small. Although this amplifier can be realized theoretically with only a single open-loop stage, a sufficiently large loop gain generally requires a cascade of three open-loop stages. As a result, pole splitting is invariably required to ensure an open-loop dominant pole response, thereby limiting the achievable closed-loop bandwidth. In addition compensation of the feedback loop may be required for acceptable closed-loop damping. The bandwidth and stability problems implicit to the use of three open-loop gain stages can be circumvented by a balanced differential realization, which requires a cascade of only two open-loop gain stages. Computationally efficient analyses are conducted by replacing the feedback subcircuit with its *y*-parameter model.
4. The series-series feedback amplifier is a transconductance signal processor in that both its input and output resistances are large. Similar to its shunt-shunt counterpart, its implementation generally requires a cascade of three open-loop gain stages. Computationally efficient analyses are conducted by replacing the feedback subcircuit with its *z*-parameter model.

The two remaining feedback circuits are dual-loop topologies that can stabilize the driving-point input and output resistances, as well as the forward gain characteristics, with respect to shifts in active element parameters. One of these latter architectures, the series-series/shunt-shunt feedback amplifier, is particularly well suited to electronic applications that require a multistage cascade.

1. The series-series/shunt-shunt feedback amplifier coalesces the series-series architecture with its shunt-shunt dual. It is particularly well suited to applications, such as wideband communication networks, which require match terminated source and load resistances. Requisite frequency compensation and broadbanding criteria mirror those incorporated in the series-series and shunt-shunt single-loop feedback topologies.
2. The series-shunt/shunt-series feedback amplifier coalesces the series-shunt architecture with its shunt-series dual. Although its input resistance can be designed to match the source resistance seen by the input port of the amplifier, and its output resistance can be matched to the load resistance driven by the amplifier, match terminated operating ($R_{in} = R_{ST} = R_{LT} = R_{out}$) is not feasible. Requisite frequency compensation and broadbanding criteria mirror those incorporated in the series-shunt and shunt-series single-loop feedback topologies.

References

[1] J. Millman and A. Grabel, *Microelectronics,* 2nd ed., New York: McGraw-Hill, 1987, chap. 12.
[2] A. B. Grebene, *Bipolar and MOS Analog Integrated Circuit Design*, New York: Wiley-Interscience, 1984, pp. 424–432.
[3] R. G. Meyer, R. Eschenbach, and R. Chin, "A wideband ultralinear amplifier from DC to 300 MHz," *IEEE J. Solid-State Circuits*, vol. SC-9, pp. 167–175, Aug. 1974.
[4] A. S. Sedra and K. C. Smith, *Microelectronic Circuits*, New York: Holt, Rinehart Winston, 1987, pp. 428–441.
[5] J. M. Early, "Effects of space-charge layer widening in junction transistors," *Proc. IRE*, vol. 46, pp. 1141–1152, Nov. 1952.
[6] A. J. Cote Jr. and J. B. Oakes, *Linear Vacuum-Tube And Transistor Circuits*, New York: McGraw-Hill, 1961, pp. 40–46.
[7] R. G. Meyer and R. A. Blauschild, "A four-terminal wideband monolithic amplifier," *IEEE J. Solid-State Circuits*, vol. SC-17, pp. 634–638, Dec. 1981.
[8] M. Ohara, Y. Akazawa, N. Ishihara, and S. Konaka, "Bipolar monolithic amplifiers for a gigabit optical repeater," *IEEE J. Solid-State Circuits*, vol. SC-19, pp. 491–497, Aug. 1985.
[9] M. J. N. Sibley, R. T. Univin, D. R. Smith, B. A. Boxall, and R. J. Hawkins, "A monolithic transimpedance preamplifier for high speed optical receivers," *British Telecommunicat. Tech. J.*, vol. 2, pp. 64–66, July 1984.
[10] J. F. Kukielka and C. P. Snapp, "Wideband monolithic cascadable feedback amplifiers using silicon bipolar technology," *IEEE Microwave Millimeter-Wave Circuits Symp. Dig.*, vol. 2, pp. 330, 331, June 1982.
[11] R. G. Meyer, M. J. Shensa, and R. Eschenbach, "Cross modulation and intermodulation in amplifiers at high frequencies," *IEEE J. Solid-State Circuits*, vol. SC-7, pp. 16–23, Feb. 1972.
[12] K. H. Chan and R. G. Meyer, "A low distortion monolithic wide-band amplifier," *IEEE J. Solid-State Circuits*, vol. SC-12, pp. 685–690, Dec. 1977.
[13] A. Arbel, "Multistage transistorized current modules," *IEEE Trans. Circuits Syst.*, vol. CT-13, pp. 302–310, Sep. 1966.
[14] A. Arbel, *Analog Signal Processing and Instrumentation*, London: Cambridge University, 1980, chap. 3.
[15] W. G. Beall, "New feedback techniques for high performance monolithic wideband amplifiers," Electron. Res. Group, University of Southern California, Tech. Memo., Jan. 1990.

13

General Feedback Theory[1]

Wai-Kai Chen
University of Illinois, Chicago

13.1 Introduction

In Chapter 11.2, we used the ideal feedback model to study the properties of feedback amplifiers. The model is useful only if we can separate a feedback amplifier into the basic amplifier $\mu(s)$ and the feedback network $\beta(s)$. The procedure is difficult and sometimes virtually impossible, because the forward path may no be strictly unilateral, the feedback path is usually bilateral, and the input and output coupling networks are often complicated. Thus, the ideal feedback model is not an adequate representation of a practical amplifier. In the remainder of this section, we shall develop Bode's feedback theory, which is applicable to the general network configuration and avoids the necessity of identifying the transfer functions $\mu(s)$ and $\beta(s)$.

Bode's feedback theory [2] is based on the concept of return difference, which is defined in terms of network determinants. We show that the return difference is a generalization of the concept of the feedback factor of the ideal feedback model, and can be measured physically from the amplifier itself. We then introduce the notion of null return difference and discuss its physical significance. Because the feedback theory will be formulated in terms of the first- and second-order cofactors of the elements of the indefinite-admittance matrix of a feedback circuit, we first review briefly the formulation of the indefinite-admittance matrix.

13.2 The Indefinite-Admittance Matrix

Figure 13.1 is an n-terminal network N composed of an arbitrary number of active and passive network elements connected in any way whatsoever. Let V_1, V_2, …, V_n be the Laplace-transformed potentials measured between terminals 1, 2, …, n and some arbitrary but unspecified reference point, and let I_1, I_2, …, I_n be the Laplace-transformed currents entering the terminals 1, 2, …, n from outside the network. The network N together with its load is linear, so the terminal current and voltages are related by the equation

[1]References for this chapter can be found on page **16**-17.

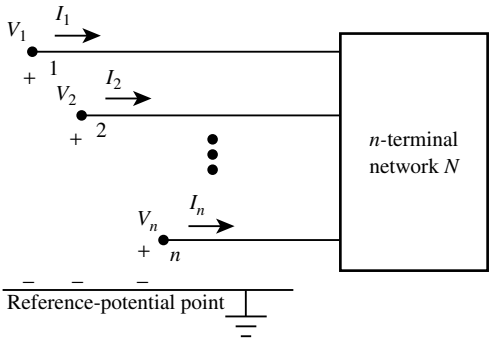


FIGURE 13.1 The general symbolic representation of an n-terminal network.

$$\begin{bmatrix} I_1 \\ I_2 \\ \vdots \\ I_n \end{bmatrix} = \begin{bmatrix} y_{11} & y_{12} & \cdots & y_{1n} \\ y_{21} & y_{22} & \cdots & y_{2n} \\ \vdots & \vdots & \vdots & \vdots \\ y_{n1} & y_{n2} & \cdots & y_{nn} \end{bmatrix} \begin{bmatrix} V_1 \\ V_2 \\ \vdots \\ V_n \end{bmatrix} + \begin{bmatrix} J_1 \\ J_2 \\ \vdots \\ J_n \end{bmatrix} \tag{13.1}$$

or more succinctly as

$$\mathbf{I}(s) = \mathbf{Y}(s)\mathbf{V}(s) + \mathbf{J}(s) \tag{13.2}$$

where J_k (k = 1,2, …, n) denotes the current flowing into the kth terminal when all terminals of N are grounded to the reference point. The coefficient matrix $\mathbf{Y}(s)$ is called the **indefinite-admittance matrix** because the reference point for the potentials is some arbitrary but unspecified point outside the network. Notice that the symbol $\mathbf{Y}(s)$ is used to denote either the admittance matrix or the indefinite-admittance matrix. This should not create any confusion because the context will tell. In the remainder of this section, we shall deal exclusively with the indefinite-admittance matrix.

We remark that the short-circuit currents J_k result from the independent sources and/or initial conditions in the interior of N. For our purposes, we shall consider all independent sources outside the network and set all initial conditions to zero. Hence, $\mathbf{J}(s)$ is considered to be zero, and (13.2) becomes

$$\mathbf{I}(s) = \mathbf{Y}(s)\mathbf{V}(s) \tag{13.3}$$

where the elements y_{ij} of $\mathbf{Y}(s)$ can be obtained as

$$y_{ij} = \left.\frac{I_i}{V_j}\right|_{v_x=0,\, x\neq j} \tag{13.4}$$

As an illustration, consider a small-signal equivalent model of a transistor in Figure 13.2. Its indefinite-admittance matrix is found to be

$$\mathbf{Y}(s) = \begin{bmatrix} g_1 + sC_1 + sC_2 & -sC_2 & -g_1 - sC_1 \\ g_m - sC_2 & g_2 + sC_2 & -g_2 - g_m \\ -g_1 - sC_1 - g_m & -g_2 & g_1 + g_2 + g_m + sC_1 \end{bmatrix} \tag{13.5}$$

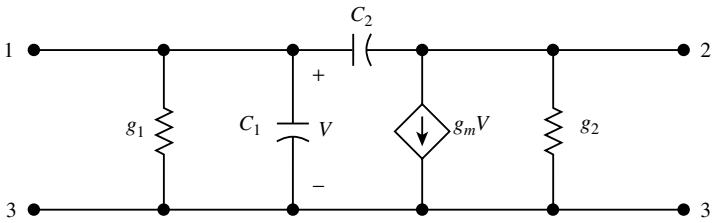


FIGURE 13.2 A small-signal equivalent network of a transistor.

Observe that the sum of elements of each row or column is equal to zero. The fact that these properties are valid in general for the indefinite-admittance matrix will now be demonstrated.

To see that the sum of the elements in each column of $\mathbf{Y}(s)$ equals zero, we add all n equations of (13.1) to yield

$$\sum_{i=1}^{n}\sum_{j=1}^{n} y_{ji}V_i = \sum_{m=1}^{n} I_m - \sum_{m=1}^{n} J_m = 0 \tag{13.6}$$

The last equation is obtained by appealing to Kirchhoff's current law (KCL) for the node corresponding to the reference point. Setting all the terminal voltages to zero except the kth one, which is nonzero, gives

$$V_k \sum_{j=1}^{n} y_{jk} = 0 \tag{13.7}$$

Because $V_k \neq 0$, it follows that the sum of the elements of each column of $\mathbf{Y}(s)$ equals zero. Thus, the indefinite-admittance matrix is always singular.

To demonstrate that each row sum of $\mathbf{Y}(s)$ is also zero, we recognize that because the point of zero potential may be chosen arbitrarily, the currents J_k and I_k remain invariant when all the terminal voltages V_k are changed by the same but arbitrary constant amount. Thus, if V_0 is an n-vector, each element of which is $v_0 \neq 0$, then

$$\mathbf{I}(s) - \mathbf{J}(s) = \mathbf{Y}(s)\left[\mathbf{V}(s) + \mathbf{V}_0\right] = \mathbf{Y}(s)\mathbf{V}(s) + \mathbf{Y}(s)\mathbf{V}_0 \tag{13.8}$$

which after invoking (13.2) yields that

$$\mathbf{Y}(s)\,\mathbf{V}_0 = \mathbf{0} \tag{13.9}$$

or

$$\sum_{j=1}^{n} y_{ij} = 0, \qquad i = 1, 2, \cdots, n \tag{13.10}$$

showing that each row sum of $\mathbf{Y}(s)$ equals zero.

Thus, if $\mathbf{Y}_{uv}$ denotes the submatrix obtained from an indefinite-admittance matrix $\mathbf{Y}(s)$ by deleting the uth row and vth column, then the **(first-order) cofactor**, denoted by the symbol Y_{uv}, of the element y_{uv} of $\mathbf{Y}(s)$, is defined by

$$Y_{uv} = (-1)^{u+v} \det \mathbf{Y}_{uv} \tag{13.11}$$

As a consequence of the zero-row-sum and zero-column-sum properties, all the cofactors of the elements of the indefinite-admittance matrix are equal. Such a matrix is also referred to as the **equicofactor matrix**. If Y_{uv} and Y_{ij} are any two cofactors of the elements of $\mathbf{Y}(s)$, then

$$Y_{uv} = Y_{ij} \tag{13.12}$$

for all *u, v, i* and *j*. For the indefinite-admittance matrix $\mathbf{Y}(s)$ of (13.5) it is straightforward to verify that all of its nine cofactors are equal to

$$Y_{uv} = s^2C_1C_2 + s\left(C_1g_2 + C_2g_1 + C_2g_2 + g_mC_2\right) + g_1g_2 \tag{13.13}$$

for $u, v = 1, 2, 3$.

Denote by $\mathbf{Y}_{rp,sq}$ the submatrix obtained from $\mathbf{Y}(s)$ by striking out rows r and s and columns p and q. Then the **second-order cofactor**, denoted by the symbol $Y_{rp,sq}$ of the elements y_{rp}, and y_{sq} of $\mathbf{Y}(s)$ is a scalar quantity defined by the relation

$$Y_{rp,sq} = \text{sgn}\left(r-s\right)\text{sgn}\left(p-q\right)\left(-1\right)^{r+p+s+q}\det\mathbf{Y}_{rp,sq} \tag{13.14}$$

where $r \neq s$ and $p \neq q$, and

$$\text{sgn } u = +1 \quad \text{if } u > 0 \tag{13.15a}$$

$$\text{sgn } u = -1 \quad \text{if } u < 0 \tag{13.15b}$$

The symbols $\mathbf{Y}_{uv}$ and Y_{uv} or $\mathbf{Y}_{rp,sq}$ and $Y_{rp,sq}$ should not create any confusion because one is in boldface whereas the other is italic. Also, for our purposes, it is convenient to define

$$Y_{rp,sq} = 0, \qquad r = s \quad \text{or} \quad p = q \tag{13.16a}$$

or

$$\text{sgn } 0 = 0 \tag{13.16b}$$

This convention will be followed throughout the remainder of this section.

As an example, consider the hybrid-pi equivalent network of a transistor in Figure 13.3. Assume that each node is an accessible terminal of a four-terminal network. Its indefinite-admittance matrix is:

$$\mathbf{Y}(s) = \begin{bmatrix} 0.02 & 0 & -0.02 & 0 \\ 0 & 5\times10^{-12}s & 0.2 - 5\times10^{-12}s & -0.2 \\ -0.02 & -5\times10^{-12}s & 0.024 + 105\times10^{-12}s & -0.004 - 10^{-10}s \\ 0 & 0 & -0.204 - 10^{-10}s & 0.204 + 10^{-10}s \end{bmatrix} \tag{13.17}$$

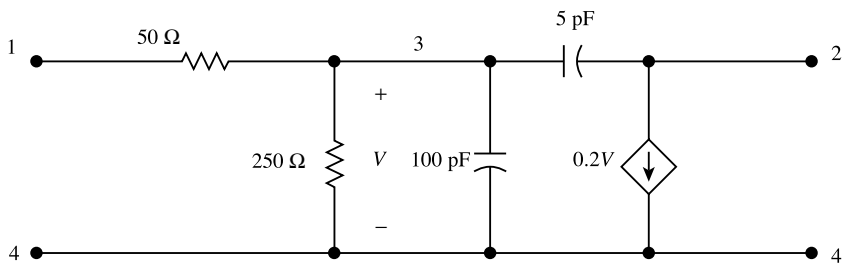


FIGURE 13.3 The hybrid-pi equivalent network of a transistor.

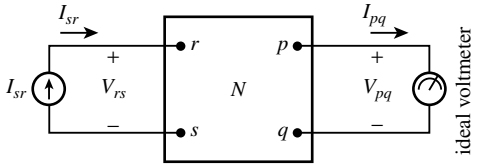


FIGURE 13.4 The symbolic representation for the measurement of the transfer impedance.

The second-order cofactor $Y_{31,42}$ and $Y_{11,34}$ of the elements of $\mathbf{Y}(s)$ of (13.17) are computed as follows:

$$Y_{31,42} = \text{sgn}(3-4)\,\text{sgn}(1-2)(-1)^{3+1+4+2}\ \det\begin{bmatrix} -0.02 & 0 \\ 0.2-5\times10^{-12}s & -0.2 \end{bmatrix} \tag{13.18a}$$

$$= 0.004$$

$$Y_{11,34} = \text{sgn}(1-3)\,\text{sgn}(1-4)(-1)^{1+1+3+4}\ \det\begin{bmatrix} 5\times10^{-12}s & 0.2-5\times10^{-12}s \\ 0 & -0.204-10^{-10}s \end{bmatrix} \tag{13.18b}$$

$$= 5\times10^{-12}s\left(0.204+10^{-10}s\right)$$

The usefulness of the indefinite-admittance matrix lies in the fact that it facilitates the computation of the driving-point or transfer functions between any pair of nodes or from any pair of nodes to any other pair. In the following, we present elegant, compact, and explicit formulas that express the network functions in terms of the ratios of the first- and/or second-order cofactors of the elements of the indefinite-admittance matrix.

Assume that a current source is connected between any two nodes r and s so that a current I_{sr} is injected into the rth node and at the same time is extracted from the sth node. Suppose that an ideal voltmeter is connected from node p to node q so that it indicates the potential rise from q to p, as depicted symbolically in Figure 13.4. Then the **transfer impedance**, denoted by the symbol $z_{rp,sq}$, between the node pairs rs and pq of the network of Figure 13. 4 is defined by the relation

$$z_{rp,sq} = \frac{V_{pq}}{I_{sr}} \tag{13.19}$$

with all initial conditions and independent sources inside N set to zero. The representation is, of course, quite general. When $r = p$ and $s = q$, the transfer impedance $z_{rp,sq}$, becomes the *driving-point impedance* $z_{rr,ss}$ between the terminal pair rs.

In Figure 13.4, set all initial conditions and independent sources in N to zero and choose terminal q to be the reference-potential point for all other terminals. In terms of (13.1), these operations are equivalent to setting $\mathbf{J} = \mathbf{0}$, $V_q = 0$, $I_x = 0$ for $x \neq r, s$ and $I_r = -I_s = I_{sr}$. Because $\mathbf{Y}(s)$ is an equicofactor matrix, the equations of (13.1) are not linearly independent and one of them is superfluous. Let us suppress the sth equation from (13.1), which then reduces to

$$\mathbf{I}_{-s} = \mathbf{Y}_{sq}\mathbf{V}_{-q} \tag{13.20}$$

where $\mathbf{I}_{-s}$ and $\mathbf{V}_{-q}$ denote the subvectors obtained from $\mathbf{I}$ and $\mathbf{V}$ of (13.3) by deleting the sth row and qth row, respectively. Applying Cramer's rule to solve for V_p yields

$$V_p = \frac{\det \tilde{\mathbf{Y}}_{sq}}{\det \mathbf{Y}_{sq}} \tag{13.21}$$

where $\tilde{\mathbf{Y}}_{sq}$ is the matrix derived from $\mathbf{Y}_{sq}$ by replacing the column corresponding to V_p by $\mathbf{I}_{-s}$. We recognize that $\mathbf{I}_{-s}$ is in the pth column if $p < q$ but in the $(p-1)$th column if $p > q$. Furthermore, the row in which I_{sr} appears is the rth row if $r < s$, but is the $(r-1)$th row if $r > s$. Thus, we obtain

$$(-1)^{s+q} \det \tilde{\mathbf{Y}}_{sq} = I_{sr} Y_{rp,sq} \tag{13.22}$$

In addition, we have

$$\det \mathbf{Y}_{sq} = (-1)^{s+q} Y_{sq} \tag{13.23}$$

Substituting these in (13.21) in conjunction with (13.19), we obtain

$$z_{rp,sq} = \frac{Y_{rp,sq}}{Y_{uv}} \tag{13.24}$$

$$z_{rr,ss} = \frac{Y_{rr,ss}}{Y_{uv}} \tag{13.25}$$

in which we have invoked the fact that $Y_{sq} = Y_{uv}$.

The *voltage gain*, denoted by $g_{rp,\,sq}$, between the node pairs rs and pq of the network of Figure 13.4 is defined by

$$g_{rp,sq} = \frac{V_{pq}}{V_{rs}} \tag{13.26}$$

again with all initial conditions and independent sources in N being set to zero. Thus, from (13.24) and (13.25) we obtain

$$g_{rp,sq} = \frac{z_{rp,sq}}{z_{rr,ss}} = \frac{Y_{rp,sq}}{Y_{rr,ss}} \tag{13.27}$$

The symbols have been chosen to help us remember. In the numerators of (13.24), (13.25), and (13.27), the order of the subscripts is as follows: r, the current injecting node; p, the voltage measurement node; s, the current extracting node; and q the voltage reference node. Nodes r and p designate the input and output transfer measurement, and nodes s and q form a sort of double datum.

As an illustration, we consider the hybrid-pi transistor equivalent network of Figure 13.3. For this transistor, suppose that we connect a 100-Ω load resistor between nodes 2 and 4, and excite the resulting circuit by a voltage source V_{14}, as depicted in Figure 13.5. To simplify our notation, let $p = 10^{-9}s$. The indefinite-admittance matrix of the amplifier is:

$$\mathbf{Y}(s) = \begin{bmatrix} 0.02 & 0 & -0.02 & 0 \\ 0 & 0.01 + 0.005p & 0.2 - 0.005p & -0.21 \\ -0.02 & -0.005p & 0.024 + 0.105p & -0.004 - 0.1p \\ 0 & -0.01 & -0.204 - 0.1p & 0.214 + 0.1p \end{bmatrix} \tag{13.28}$$

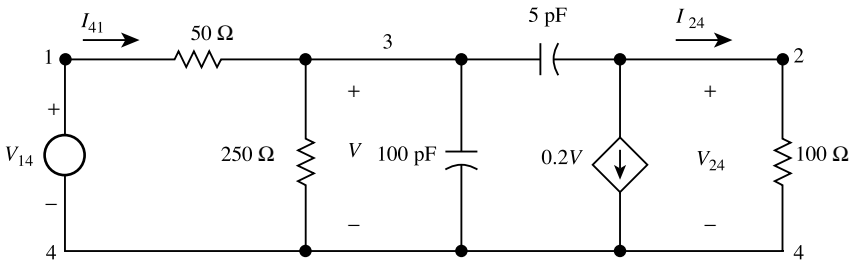


FIGURE 13.5 A transistor amplifier used to illustrate the computation of $g_{rp,sq}$.

To compute the voltage gain $g_{12,44}$, we appeal to (13.27) and obtain

$$g_{12,44} = \frac{V_{24}}{V_{14}} = \frac{Y_{12,44}}{Y_{11,44}} = \frac{p-40}{5p^2+21.7p+2.4} \tag{13.29}$$

The input impedance facing the voltage source V_{14} is determined by

$$z_{11,44} = \frac{V_{14}}{I_{41}} = \frac{Y_{11,44}}{Y_{uv}} = \frac{Y_{11,44}}{Y_{44}} = \frac{50p^2+217p+24}{p^2+4.14p+0.08} \tag{13.30}$$

To compute the current gain defined as the ratio of the current I_{24} in the 100-Ω resistor to the input current I_{41}, we apply (13.24) and obtain

$$\frac{I_{24}}{I_{41}} = 0.01\frac{V_{24}}{I_{41}} = 0.01z_{12,44} = 0.01\frac{Y_{12,44}}{Y_{44}} = \frac{0.1p-4}{p^2+4.14p+0.08} \tag{13.31}$$

Finally, to compute the transfer admittance defined as the ratio of the load current I_{24} to the input voltage V_{14}, we appeal to (13.27) and obtain

$$\frac{I_{24}}{V_{14}} = 0.01\frac{V_{24}}{V_{14}} = 0.01g_{12,44} = 0.01\frac{Y_{12,44}}{Y_{11,44}} = \frac{p-40}{500p^2+2170p+240} \tag{13.32}$$

13.3 The Return Difference

In the study of feedback amplifier response, we are usually interested in how a particular element of the amplifier affects that response. This element is either crucial in terms of its effect on the entire system or of primary concern to the designer. It may be the transfer function of an active device, the gain of an amplifier, or the immittance of a one-port network. For our purposes, we assume that this element x is the controlling parameter of a voltage-controlled current source defined by the equation

$$I = xV \tag{13.33}$$

To focus our attention on the element x, Figure 13.6 is the general configuration of a feedback amplifier in which the controlled source is brought out as a two-port network connected to a general four-port network, along with the input source combination of I_s and admittance Y_1 and the load admittance Y_2.

We remark that the two-port representation of a controlled source (13.33) is quite general. It includes the special situation where a one-port element is characterized by its immittance. In this case, the controlling voltage V is the terminal voltage of the controlled current source I, and x become the one-port admittance.

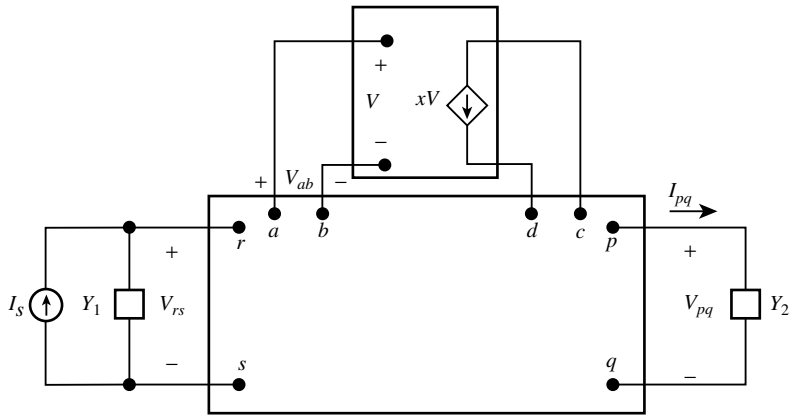


FIGURE 13.6 The general configuration of a feedback amplifier.

The *return difference* $F(x)$ of a feedback amplifier with respect to an element x is defined as the ratio of the two functional values assumed by the first-order cofactor of an element of its indefinite-admittance matrix under the condition that the element x assumes its nominal value and the condition that the element x assumes the value zero. To emphasize the importance of the feedback element x, we express the indefinite-admittance matrix $\mathbf{Y}$ of the amplifier as a function of x, even though it is also a function of the complex-frequency variable s, and write $\mathbf{Y} = \mathbf{Y}(x)$. Then, we have [3]

$$F(x) \equiv \frac{Y_{uv}(x)}{Y_{uv}(0)} \tag{13.34}$$

where

$$Y_{uv}(0) = Y_{uv}(x)\big|_{x=0} \tag{13.35}$$

The physical significance of the return difference will now be considered. In the network of Figure 13.6, the input, the output, the controlling branch, and the controlled source are labeled as indicated. Then, the element x enters the indefinite-admittance matrix $\mathbf{Y}(x)$ in a rectangular pattern as shown next:

$$\mathbf{Y}(x) = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{bmatrix} & & & \\ & & & \\ x & -x & & \\ -x & x & & \end{bmatrix} \end{array} \tag{13.36}$$

If in Figure 13.6 we replace the controlled current source xV by an independent current source of x A and set the excitation current source I_s to zero, the indefinite-admittance matrix of the resulting network is simply $\mathbf{Y}(0)$. By appealing to (13.24), the new voltage V'_{ab} appearing at terminals a and b of the controlling branch is:

$$V'_{ab} = x\frac{Y_{da,cb}(0)}{Y_{uv}(0)} = -x\frac{Y_{ca,db}(0)}{Y_{uv}(0)} \tag{13.37}$$

Notice that the current injecting point is terminal d, not c.

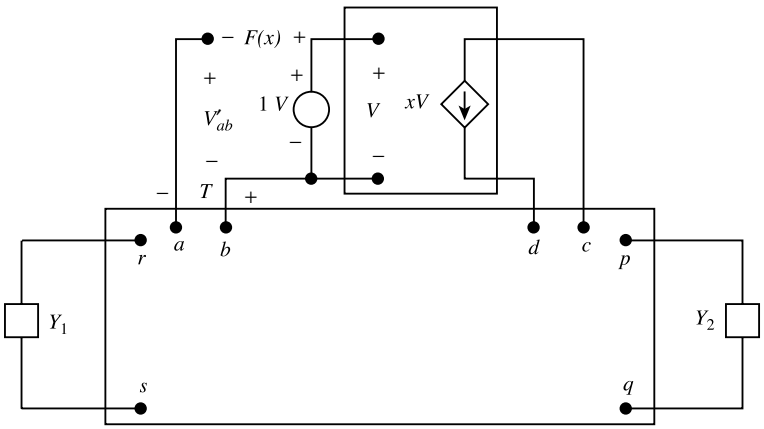


FIGURE 13.7 The physical interpretation of the return difference with respect to the controlling parameter of a voltage-controlled current source.

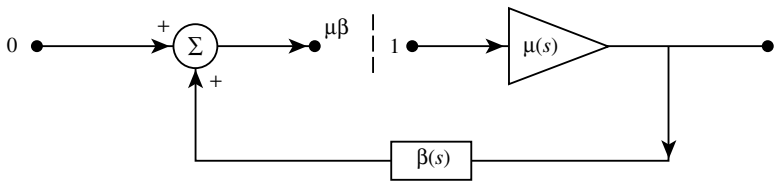


FIGURE 13.8 The physical interpretation of the loop transmission.

The preceding operation of replacing the controlled current source by an independent current source and setting the excitation I_s to zero can be represented symbolically as in Figure 13.7. Observe that the controlling branch is broken off as marked and a 1-V voltage source is applied to the right of the breaking mark. This 1-V sinusoidal voltage of a fixed angular frequency produces a current of x A at the controlled current source. The voltage appearing at the left of the breaking mark caused by this 1-V excitation is then V'_{ab} as indicated. This returned voltage V'_{ab} has the same physical significance as the loop transmission $\mu\beta$ defined for the ideal feedback model in Chapter 11. To see this, we set the input excitation to the ideal feedback model to zero, break the forward path, and apply a unit input to the right of the break, as depicted in Figure 13.8. The signal appearing at the left of the break is precisely the loop transmission.

For this reason, we introduce the concept of **return ratio** T, which is defined as the negative of the voltage appearing at the controlling branch when the controlled current source is replaced by an independent current source of x A and the input excitation is set to zero. Thus, the return ratio T is simply the negative of the returned voltage V'_{ab}, or $T = -V'_{ab}$. With this in mind, we next compute the difference between the 1-V excitation and the returned voltage V'_{ab} obtaining

$$\begin{aligned} 1 - V'_{ab} &= 1 + x\frac{Y_{ca,db}}{Y_{uv}(0)} = \frac{Y_{uv}(0) + xY_{ca,db}}{Y_{uv}(0)} = \frac{Y_{db}(0) + xY_{ca,db}}{Y_{db}(0)} \\ &= \frac{Y_{db}(x)}{Y_{db}(0)} = \frac{Y_{uv}(x)}{Y_{uv}(0)} = F(x) \end{aligned} \tag{13.38}$$

in which we have invoked the indentities $Y_{uv} = Y_{ij}$ and

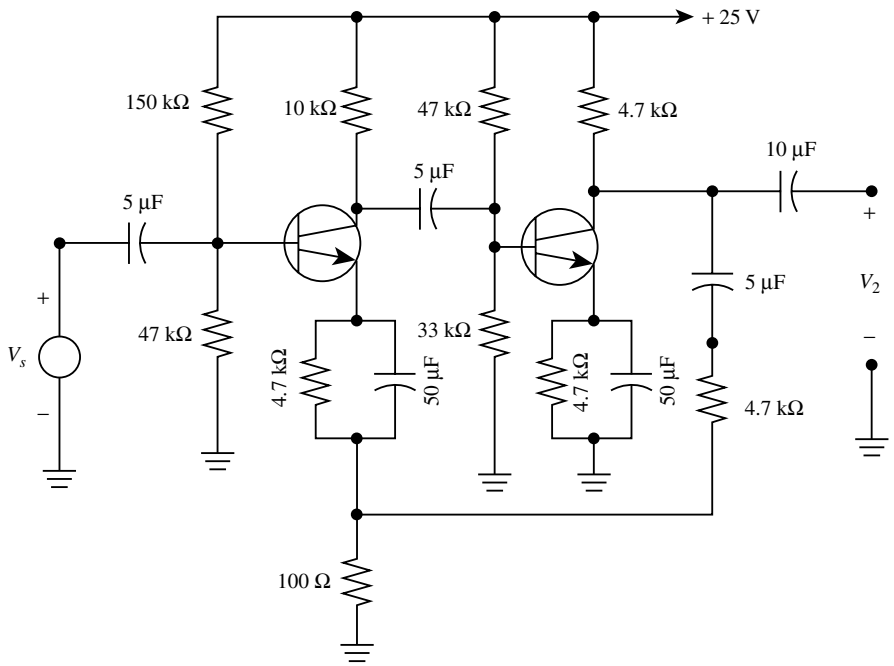


FIGURE 13.9 A voltage-series feedback amplifier together with its biasing and coupling circuitry.

$$Y_{db}(x) = Y_{db}(0) + xY_{ca,db} \tag{13.39}$$

We remark that we write $Y_{ca,db}(x)$ as $Y_{ca,db}$ because it is independent of x. In other words, the return difference $F(x)$ is simply the difference of the 1-V excitation and the returned voltage V'_{ab} as illustrated in Figure 13.7, and hence its name. Because

$$F(x) = 1 + T = 1 - \mu\beta \tag{13.40}$$

we conclude that the return difference has the same physical significance as the feedback factor of the ideal feedback model. The significance of the previous physical interpretations is that it permits us to determine the return ratio T or $-\mu\beta$ by measurement. Once the return ratio is measured, the other quantities such as return difference and loop transmission are known.

To illustrate, consider the voltage-series or the series-parallel feedback amplifier of Figure 13.9. Assume that the two transistors are identical with the following hybrid parameters:

$$h_{ie} = 1.1\ \text{k}\Omega, \qquad h_{fe} = 50, \qquad h_{re} = h_{oe} = 0 \tag{13.41}$$

After the biasing and coupling circuitry have been removed, the equivalent network is presented in Figure 13.10. The effective load of the first transistor is composed of the parallel combination of the 10, 33, 47, and 1.1-kΩ resistors. The effect of the 150- and 47-kΩ resistors can be ignored; they are included in the equivalent network to show their insignificance in the computation.

To simplify our notation, let

$$\tilde{\alpha}_k = \alpha_k \times 10^{-4} = \frac{h_{fe}}{h_{ie}} = 455 \times 10^{-4}, \qquad k = 1, 2 \tag{13.42}$$

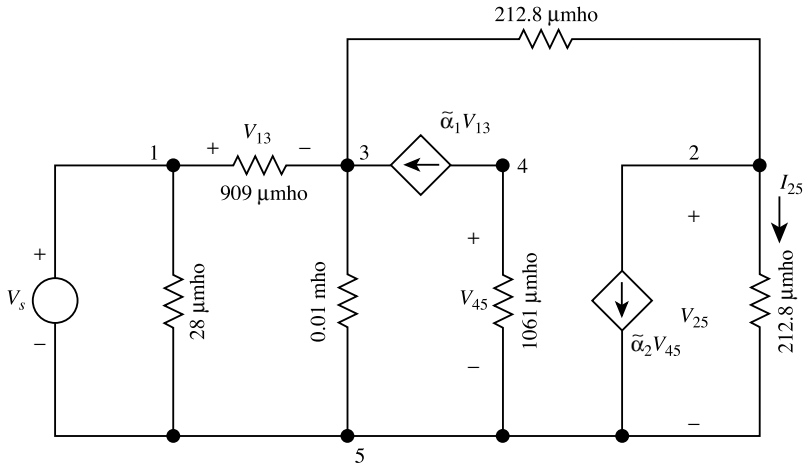


FIGURE 13.10 An equivalent network of the feedback amplifier of Figure 13.9.

The subscript k is used to distinguish the transconductances of the first and the second transistors. The indefinite-admittance matrix of the feedback amplifier of Figure 13.9 is:

$$\mathbf{Y} = 10^{-4}\begin{bmatrix} 9.37 & 0 & -9.09 & 0 & -0.28 \\ 0 & 4.256 & -2.128 & \alpha_2 & -2.128-\alpha_2 \\ -9.09-\alpha_1 & -2.128 & 111.218+\alpha_1 & 0 & -100 \\ \alpha_1 & 0 & -\alpha_1 & 10.61 & -10.61 \\ -0.28 & -2.128 & -100 & -10.61-\alpha_2 & 113.018+\alpha_2 \end{bmatrix} \tag{13.43}$$

By applying (13.27), the amplifier voltage gain is computed as

$$g_{12,25} = \frac{V_{25}}{V_s} = \frac{V_{12,25}}{V_{11,25}} = \frac{211.54\times10^{-7}}{4.66\times10^{-7}} = 45.39 \tag{13.44}$$

To calculate the return differences with respect to the transconductances $\tilde{\alpha}_k$ of the transistors, we short-circuit the voltage source V_s. The resulting indefinite-admittance matrix is obtained from (13.43) by adding the first row to the fifth row and the first column to the fifth column and then deleting the first row and column. Its first-order cofactor is simply $Y_{11,55}$. Thus, the return differences with respect to $\tilde{\alpha}_k$ are:

$$F(\tilde{\alpha}_1) = \frac{Y_{11,55}(\tilde{\alpha}_1)}{Y_{11,55}(0)} = \frac{466.1\times10^{-9}}{4.97\times10^{-9}} = 93.70 \tag{13.45a}$$

$$F(\tilde{\alpha}_2) = \frac{Y_{11,55}(\tilde{\alpha}_2)}{Y_{11,55}(0)} = \frac{466.1\times10^{-9}}{25.52\times10^{-9}} = 18.26 \tag{13.45b}$$

13.4 The Null Return Difference

In this section, we introduce the notion of null return difference, which is found to be very useful in measurement situations and in the computation of the sensitivity for the feedback amplifiers.

The **null return difference** $\hat{F}(x)$ of a feedback amplifier with respect to an element x is defined to be the ratio of the two functional values assumed by the second-order cofactor $Y_{rp,sq}$ of the elements of its indefinite-admittance matrix **Y** under the condition that the element x assumes its nominal value and the condition that the element x assumes the value zero where r and s are input terminals, and p and q are the output terminals of the amplifier, or

$$\hat{F}(x) = \frac{Y_{rp,sq}(x)}{Y_{rp,sq}(0)} \tag{13.46}$$

Likewise, the **null return ratio** $\hat{T}$, with respect to a voltage-controlled current source $I = xV$, is the negative of the voltage appearing at the controlling branch when the controlled current source is replaced by an independent current source of x A and when the input excitation is adjusted so that the output of the amplifier is identically zero.

Now, we demonstrate that the null return difference is simply the return difference in the network under the situation that the input excitation I_s has been adjusted so that the output is identically zero. In the network of Figure 13.6, suppose that we replace the controlled current source by an independent current source of x A. Then by applying formula (13.24) and the superposition principle, the output current I_{pq} at the load is:

$$I_{pq} = Y_2\left[I_s\frac{Y_{rp,sq}(0)}{Y_{uv}(0)} + x\frac{Y_{dp,cq}(0)}{Y_{uv}(0)}\right] \tag{13.47}$$

Setting $I_{pq} = 0$ or $V_{pq} = 0$ yields

$$I_s \equiv I_0 = -x\left[\frac{Y_{dp,cq}(0)}{Y_{rp,sq}(0)}\right] \tag{13.48}$$

in which $Y_{dp,cq}$ is independent of x. This adjustment is possible only if a direct transmission occurs from the input to the output when x is set to zero. Thus, in the network of Figure 13.7, if we connect an independent current source of strength I_0 at its input port, the voltage V'_{ab} is the negative of the null return ratio $\hat{T}$. Using (13.24), we obtain [4]

$$\begin{aligned}\hat{T} = -V'_{ab} &= -x\frac{Y_{da,cb}(0)}{Y_{uv}(0)} - I_0\frac{Y_{ra,sb}(0)}{Y_{uv}(0)} \\ &= -\frac{x\left[Y_{da,cb}(0)Y_{rp,sq}(0) - Y_{ra,sb}(0)Y_{dp,cq}(0)\right]}{Y_{uv}(0)Y_{rp,sq}(0)} \\ &= \frac{x\dot{Y}_{rp,sq}}{Y_{rp,sq}(0)} = \frac{Y_{rp,sq}(x)}{Y_{rp,sq}(0)} - 1\end{aligned} \tag{13.49}$$

where

$$\dot{Y}_{rp,sq} \equiv \frac{dY_{rp,sq}(x)}{dx} \tag{13.50}$$

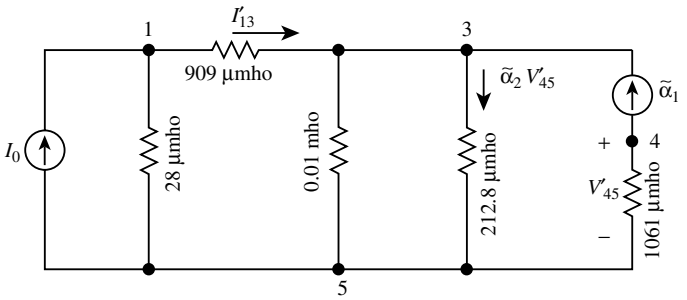


FIGURE 13.11 The network used to compute the null return difference $\hat{F}(\tilde{\alpha}_1)$ by its physical interpretation.

This leads to

$$\hat{F}(x) = 1 + \hat{T} = 1 - V'_{ab} \tag{13.51}$$

which demonstrates that the null return difference $\hat{F}(x)$ is simply the difference of the 1-V excitation applied to the right of the breaking mark of the broken controlling branch of the controlled source and the returned voltage V'_{ab} appearing at the left of the breaking mark under the situation that the input signal I_s is adjusted so that the output is identically zero.

As an illustration, consider the voltage-series feedback amplifier of Figure 13.9, an equivalent network of which is presented in Figure 13.10. Using the indefinite-admittance matrix of (13.43) in conjunction with (13.42), the null return differences with respect to $\hat{\alpha}_k$ are:

$$\hat{F}(\tilde{\alpha}_1) = \frac{Y_{12,55}(\tilde{\alpha}_1)}{Y_{12,55}(0)} = \frac{211.54 \times 10^{-7}}{205.24 \times 10^{-12}} = 103.07 \times 10^3 \tag{13.52a}$$

$$\hat{F}(\tilde{\alpha}_2) = \frac{Y_{12,55}(\tilde{\alpha}_2)}{Y_{12,55}(0)} = \frac{211.54 \times 10^{-7}}{104.79 \times 10^{-10}} = 2018.70 \tag{13.52b}$$

Alternatively, $\hat{F}(\tilde{\alpha}_1)$ can be computed by using its physical interpretation as follows. Replace the controlled source $\tilde{\alpha}_1 V_{13}$ in Figure 13.10 by an independent current source of $\tilde{\alpha}_1$ A. We then adjust the voltage source V_s so that the output current I_{25} is identically zero. Let I_0 be the input current resulting from this source. The corresponding network is presented in Figure 13.11. From this network, we obtain

$$\hat{F}(\tilde{\alpha}_1) = 1 + \hat{T} = 1 - V'_{13} = 1 - \frac{100V'_{35} + \alpha_2 V'_{45} - \alpha_1}{9.09} = 103.07 \times 10^3 \tag{13.53}$$

Likewise, we can use the same procedure to compute the return difference $\hat{F}(\tilde{\alpha}_2)$.

14

The Network Functions and Feedback[1]

Wai-Kai Chen
University of Illinois, Chicago

We now study the effects of feedback on amplifier impedance and gain and obtain some useful relations among the return difference, the null return difference, and impedance functions in general.

Refer to the general feedback configuration of Figure 13.6. Let w be a transfer function. As before, to emphasize the importance of the feedback element x, we write $w = w(x)$. To be definite, let $w(x)$ for the time being be the current gain between the output and input ports. Then, from (13.24) we obtain

$$w(x) = \frac{I_{pq}}{I_s} = \frac{Y_2 V_{pq}}{I_s} = \frac{Y_{rp,sq}(x)}{Y_{uv}(x)} Y_2 \tag{14.1}$$

yielding

$$\frac{w(x)}{w(0)} = \frac{Y_{rp,sq}(x)}{Y_{uv}(x)} \frac{Y_{uv}(0)}{Y_{rp,sq}(0)} = \frac{\hat{F}(x)}{F(x)} \tag{14.2}$$

provided that $w(0) \neq 0$. This gives a very useful formula for computing the current gain:

$$w(x) = w(0) \frac{\hat{F}(x)}{F(x)} \tag{14.3}$$

Equation (14.3) remains valid if $w(x)$ represents the transfer impedance $z_{rp,sq} = V_{pq}/I_S$ instead of the current gain.

14.1 Blackman's Formula

In particular, when $r = p$ and $s = q$, $w(x)$ represents the driving-point impedance $z_{rr,ss}(x)$ looking into the terminals r and s, and we have a somewhat different interpretation. In this case, $F(x)$ is the return difference with respect to the element x under the condition $I_s = 0$. Thus, $F(x)$ is the return difference for the situation when the port where the input impedance is defined is left open without a source and

[1]References for this chapter can be found on page **16**-17.

we write $F(x) = F$(input open circuited). Likewise, from Figure 13.6, $\hat{F}(x)$ is the return difference with respect to x for the input excitation I_s and output response V_{rs} under the condition I_s is adjusted so that V_{rs} is identically zero. Thus, $\hat{F}(x)$ is the return difference for the situation when the port where the input impedance is defined is short circuited, and we write $\hat{F}(x) = F$(input short-circuited). Consequently, the input impedance $Z(x)$ looking into a terminal pair can be conveniently expressed as

$$Z(x) = Z(0)\frac{F(\text{input short circuited})}{F(\text{input open circuited})} \tag{14.4}$$

This is the well-known **Blackman's formula** for computing an active impedance. The formula is extremely useful because the right-hand side can usually be determined rather easily. If x represents the controlling parameter of a controlled source in a single-loop feedback amplifier, then setting $x = 0$ opens the feedback loop and $Z(0)$ is simply a passive impedance. The return difference for x when the input port is short circuited or open circuited is relatively simple to compute because shorting out or opening a terminal pair frequently breaks the feedback loop. In addition, Blackman's formula can be used to determine the return difference by measurements. Because it involves two return differences, only one of them can be identified and the other must be known in advance. In the case of a single-loop feedback amplifier, it is usually possible to choose a terminal pair so that either the numerator or the denominator on the right-hand side of (14.4) is unity. If F(input short circuited) = 1, F(input open circuited) becomes the return difference under normal operating condition and we have

$$F(x) = \frac{Z(0)}{Z(x)} \tag{14.5}$$

On the other hand, if F(input open-circuited) = 1, F(input short-circuited) becomes the return difference under normal operating condition and we obtain

$$F(x) = \frac{Z(x)}{Z(0)} \tag{14.6}$$

Example 1. The network of Figure 14.1 is a general active RC one-port realization of a rational impedance. We use Blackman's formula to verify that its input admittance is given by

$$Y = 1 + \frac{Z_3 - Z_4}{Z_1 - Z_2} \tag{14.7}$$

Appealing to (14.4), the input admittance written as $Y = Y(x)$ can be written as

$$Y(x) = Y(0)\frac{F(\text{input open circuited})}{F(\text{input short circuited})} \tag{14.8}$$

where $x = 2/Z_3$. By setting x to zero, the network used to compute $Y(0)$ is shown in Figure 14.2. Its input admittance is:

$$Y(0) = \frac{Z_1 + Z_2 + Z_3 + Z_4 + 2}{Z_1 + Z_2} \tag{14.9}$$

When the input port is open-circuited, the network of Figure 14.1 degenerates to that depicted in Figure 14.3. The return difference with respect to x is:

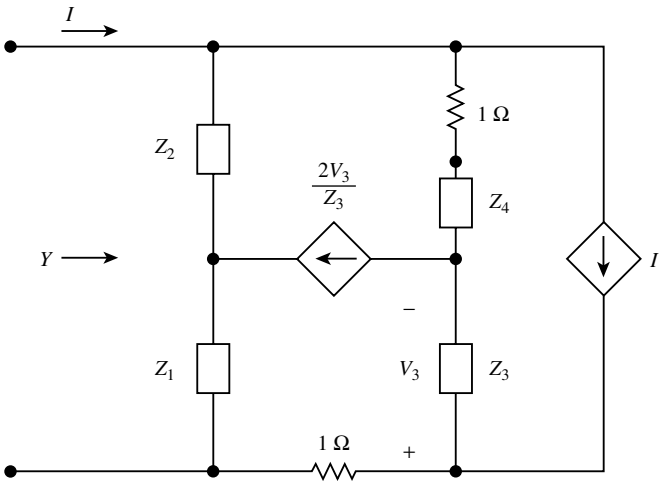


FIGURE 14.1 A general active RC one-port realization of a rational function.

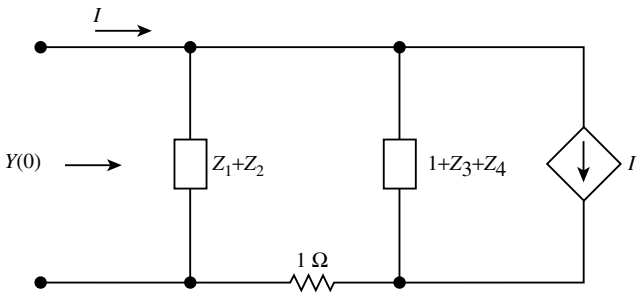


FIGURE 14.2 The network used to compute $Y(0)$.

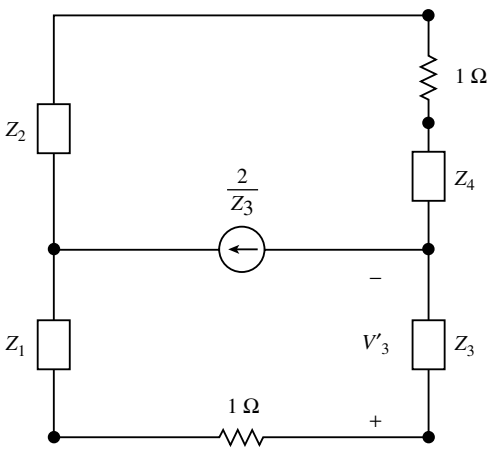


FIGURE 14.3 The network used to compute F(input open-circuited).

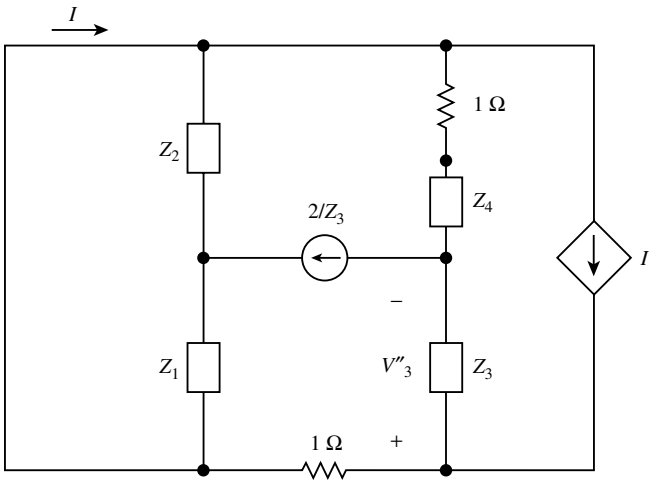


FIGURE 14.4 The network used to compute F(input short-circuited).

$$F\left(\text{input open-circuited}\right)=1-V_3'=\frac{Z_1+Z_3-Z_2-Z_4}{2+Z_1+Z_2+Z_3+Z_4} \tag{14.10}$$

where the returned voltage V_3' at the controlling branch is given by

$$V_3'=\frac{2\left(1+Z_2+Z_4\right)}{2+Z_1+Z_2+Z_3+Z_4} \tag{14.11}$$

To compute the return difference when the input port is short circuited, we use the network of Figure 14.4 and obtain

$$F\left(\text{input short-circuited}\right)=1-V_3''=\frac{Z_1-Z_2}{Z_1+Z_2} \tag{14.12}$$

where the return voltage V_3'' at the controlling branch is found to be

$$V_3''=\frac{2Z_2}{Z_1+Z_2} \tag{14.13}$$

Substituting (14.9), (14.10), and (14.12) in (14.8) yields the desired result.

$$Y=1+\frac{Z_3-Z_4}{Z_1-Z_2} \tag{14.14}$$

To determine the effect of feedback on the input and output impedances, we choose the series-parallel feedback configuration of Figure 14.5. By shorting the terminals of Y_2, we interrupt the feedback loop, therefore, formula (14.5) applies and the output impedance across the load admittance Y_2 becomes

$$Z_{\text{out}}\left(x\right)=\frac{Z_{\text{out}}\left(0\right)}{F\left(x\right)} \tag{14.15}$$

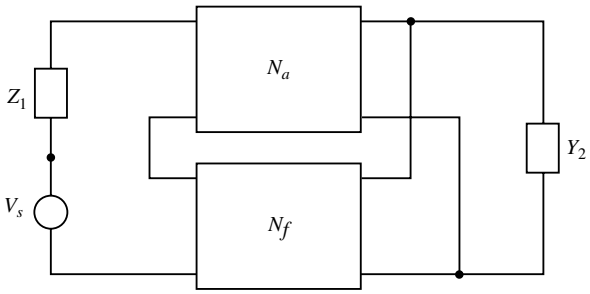


FIGURE 14.5 The series-parallel feedback configuration.

demonstrating that the impedance measured across the path of the feedback is reduced by the factor that is the normal value of the return difference with respect to the element x, where x is an arbitrary element of interest. For the input impedance of the amplifier looking into the voltage source V_s of Figure 14.5, by open circuiting or removing the voltage source V_s, we break the feedback loop. Thus, formula (14.6) applies and the input impedance becomes

$$Z_{\text{in}}(x) = F(x)Z_{\text{in}}(0) \tag{14.16}$$

meaning that the impedance measured in series lines is increased by the same factor $F(x)$. Similar conclusions can be reached for other types of configurations discussed in Chapter 12 by applying Blackman's formula.

Again, refer to the general feedback configuration of Figure 13.6 If $w(x)$ represents the voltage gain V_{pq}/V_{rs} or the transfer admittance I_{pq}/V_{rs}. Then, from (13.27) we can write

$$\frac{w(x)}{w(0)} = \frac{Y_{rp,sq}(x)}{Y_{rp,sq}(0)} \frac{Y_{rr,ss}(0)}{Y_{rr,ss}(x)} \tag{14.17}$$

The first term in the product on the right-hand side is the null return difference $\hat{F}(x)$ with respect to x for the input terminals r and s and output terminals p and q. The second term is the reciprocal of the null return difference with respect to x for the same input and output port at terminals r and s. This reciprocal can then be interpreted as the return difference with respect to x when the input port of the amplifier is short circuited. Thus, the voltage gain or the transfer admittance can be expressed as

$$w(x) = w(0)\frac{\hat{F}(x)}{F(\text{input short-circuited})} \tag{14.18}$$

Finally, if w (x)denotes the short circuit current gain I_{pq}/I_s as Y_2 approaches infinity, we obtain

$$\frac{w(x)}{w(0)} = \frac{Y_{rp,sq}(x)}{Y_{rp,sq}(0)} \frac{Y_{pp,qq}(0)}{Y_{pp,qq}(x)} \tag{14.19}$$

The second term in the product on the right-hand side is the reciprocal of the return difference with respect to x when the output port of the amplifier is short-circuited, giving a formula for the short circuit current gain as

$$w(x) = w(0)\frac{\hat{F}(x)}{F(\text{output short-circuited})} \tag{14.20}$$

Again, consider the voltage-series or series-parallel feedback amplifier of Figure 13.9 an equivalent network of which is given in Figure 13.10. The return differences $F(\tilde{\alpha}_k)$, the null return differences $\hat{F}(\tilde{\alpha}_k)$ and the voltage gain w were computed earlier in (13.45), (13.52), and (13.44), and are repeated next:

$$F(\tilde{\alpha}_1) = 93.70, \qquad F(\tilde{\alpha}_2) = 18.26 \tag{14.21a}$$

$$\hat{F}(\tilde{\alpha}_1) = 103.07 \times 10^3, \quad \hat{F}(\tilde{\alpha}_2) = 2018.70 \tag{14.21b}$$

$$w = \frac{V_{25}}{V_s} = w(\tilde{\alpha}_1) = w(\tilde{\alpha}_2) = 45.39 \tag{14.21c}$$

We apply (14.18) to calculate the voltage gain w, as follows:

$$w(\tilde{\alpha}_1) = w(0)\frac{\hat{F}(\tilde{\alpha}_1)}{F(\text{input short-circuited})} = 0.04126\,\frac{103.07 \times 10^3}{93.699} = 45.39 \tag{14.22}$$

where

$$w(0) = \left.\frac{Y_{12,55}(\tilde{\alpha}_1)}{Y_{11,55}(\tilde{\alpha}_1)}\right|_{\tilde{\alpha}_1=0} = \frac{205.24 \times 10^{-12}}{497.41 \times 10^{-11}} = 0.04126 \tag{14.23a}$$

$$F(\text{input short-circuited}) = \frac{Y_{11,55}(\tilde{\alpha}_1)}{Y_{11,55}(0)} = \frac{466.07 \times 10^{-9}}{4.9741 \times 10^{-9}} = 93.699 \tag{14.23b}$$

and

$$w(\tilde{\alpha}_2) = w(0)\,\frac{\hat{F}(\tilde{\alpha}_2)}{F(\text{input short-circuited})} = 0.41058\frac{2018.70}{18.26} = 45.39 \tag{14.24}$$

where

$$w(0) = \left.\frac{Y_{12,55}(\tilde{\alpha}_2)}{Y_{11,55}(\tilde{\alpha}_2)}\right|_{\tilde{\alpha}_2=0} = \frac{104.79 \times 10^{-10}}{255.22 \times 10^{-10}} = 0.41058 \tag{14.25a}$$

$$F(\text{input short-circuited}) = \frac{Y_{11,55}(\tilde{\alpha}_2)}{Y_{11,55}(0)} = \frac{466.07 \times 10^{-9}}{25.52 \times 10^{-9}} = 18.26 \tag{14.25b}$$

14.2 The Sensitivity Function

One of the most important effects of negative feedback is its ability to make an amplifier less sensitive to the variations of its parameters because of aging, temperature variations, or other environment changes. A useful quantitative measure for the degree of dependence of an amplifier on a particular parameter is known as the sensitivity. The **sensitivity function**, written as $\mathcal{S}(x)$, for a given transfer function with respect to an element x is defined as the ratio of the fractional change in a transfer function to the fractional change in x for the situation when all changes concerned are differentially small. Thus, if $w(x)$ is the transfer function, the sensitivity function can be written as

$$\mathcal{S}(x) = \lim_{\Delta x \to 0} \frac{\Delta w/w}{\Delta x/x} = \frac{x}{w}\,\frac{\partial w}{\partial x} = x\frac{\partial \ln w}{\partial x} \tag{14.26}$$

Refer to the general feedback configuration of Figure 13.6, and let $w(x)$ represent either the current gain I_{pq}/I_s or the transfer impedance V_{pq}/I_s for the time being. Then, we obtain from (13.24)

$$w(x) = Y_2 \frac{Y_{rp,sq}(x)}{Y_{uv}(x)} \quad \text{or} \quad \frac{Y_{rp,sq}(x)}{Y_{uv}(x)} \tag{14.27}$$

As before, we write

$$\dot{Y}_{uv}(x) = \frac{\partial Y_{uv}(x)}{\partial x} \tag{14.28a}$$

$$\dot{Y}_{rp,sq}(x) = \frac{\partial Y_{rp,sq}(x)}{\partial x} \tag{14.28b}$$

obtaining

$$Y_{uv}(x) = Y_{uv}(0) + x\dot{Y}_{uv}(x) \tag{14.29a}$$

$$Y_{rp,sq}(x) = Y_{rp,sq}(0) + x\dot{Y}_{rp,sq}(x) \tag{14.29b}$$

Substituting (14.27) in (14.26), in conjunction with (14.29), yields

$$\begin{aligned} \mathcal{S}(x) &= x\frac{\dot{Y}_{rp,sq}(x)}{Y_{rp,sq}(x)} - x\frac{\dot{Y}_{uv}(x)}{Y_{uv}(x)} = \frac{Y_{rp,sq}(x) - Y_{rp,sq}(0)}{Y_{rp,sq}(x)} - \frac{Y_{uv}(x) - Y_{uv}(0)}{Y_{uv}(x)} \\ &= \frac{Y_{uv}(0)}{Y_{uv}(x)} - \frac{Y_{rp,sq}(0)}{Y_{rp,sq}(x)} = \frac{1}{F(x)} - \frac{1}{\hat{F}(x)} \end{aligned} \tag{14.30}$$

Combining this with (14.3), we obtain

$$\mathcal{S}(x) = \frac{1}{F(x)}\left[1 - \frac{w(0)}{w(x)}\right] \tag{14.31}$$

Observe that if $w(0) = 0$, (14.31) becomes

$$\mathcal{S}(x) = \frac{1}{F(x)} \tag{14.32}$$

meaning that sensitivity is equal to the reciprocal of the return difference. For the ideal feedback model, the feedback path is unilateral. Hence, $w(0) = 0$ and

$$\mathcal{S} = \frac{1}{F} = \frac{1}{1+T} = \frac{1}{1-\mu\beta} \tag{14.33}$$

For a practical amplifier, $w(0)$ is usually very much smaller than $w(x)$ in the passband, and $F \approx 1/\mathcal{S}$ may be used as a good estimate of the reciprocal of the sensitivity in the same frequency band. A single-loop feedback amplifier composed of a cascade of common-emitter stages with a passive network providing the desired feedback fulfills this requirements. If in such a structure any one of the transistors fails, the forward transmission is nearly zero and $w(0)$ is practically zero. Our conclusion is that if the failure of any element will interrupt the transmission through the amplifier as a whole to nearly zero, the sensitivity is approximately equal to the reciprocal of the return difference with respect to that element.

In the case of driving-point impedance, $w(0)$ is not usually smaller than $w(x)$, and the reciprocity relation is not generally valid.

Now assume that $w(x)$ represents the voltage gain. Substituting (14.27) in (14.26) results in

$$\begin{aligned}\mathcal{S}(x) &= x\frac{\dot{Y}_{rp,sq}(x)}{Y_{rp,sq}(x)} - x\frac{\dot{Y}_{rr,ss}(x)}{Y_{rr,ss}(x)} = \frac{Y_{rp,sq}(x)-Y_{rp,sq}(0)}{Y_{rp,sq}(x)} - \frac{Y_{rr,ss}(x)-Y_{rr,ss}(0)}{Y_{rr,ss}(x)} \\ &= \frac{Y_{rr,ss}(0)}{Y_{rr,ss}(x)} - \frac{Y_{rp,sq}(0)}{Y_{rp,sq}(x)} = \frac{1}{F(\text{input short-circuited})} - \frac{1}{\hat{F}(x)}\end{aligned} \tag{14.34}$$

Combining this with (14.18) gives

$$\mathcal{S}(x) = \frac{1}{F(\text{input short-circuited})}\left[1-\frac{w(0)}{w(x)}\right] \tag{14.35}$$

Finally, if $w(x)$ denotes the short circuit current gain I_{pq}/I_s as Y_2 approaches infinity, the sensitivity function can be written as

$$\mathcal{S}(x) = \frac{Y_{pp,qq}(0)}{Y_{pp,qq}(x)} - \frac{Y_{rp,sq}(0)}{Y_{rp,sq}(x)} = \frac{1}{F(\text{output short-circuited})} - \frac{1}{\hat{F}(x)} \tag{14.36}$$

which when combined with (14.20) yields

$$\mathcal{S}(x) = \frac{1}{F(\text{output short-circuited})}\left[1-\frac{w(0)}{w(x)}\right] \tag{14.37}$$

We remark that formulas (14.31), (14.35), and (14.39) are quite similar. If the return difference $F(x)$ is interpreted properly, they can all be represented by the single relation (14.31). As before, if $w(0) = 0$, the sensitivity for the voltage gain function is equal to the reciprocal of the return difference under the situation that the input port of the amplifier is short-circuited, whereas the sensitivity for the short circuit current gain is the reciprocal of the return difference when the output port is short-circuited.

Example 2. The network of Figure 14.6 is a common-emitter transistor amplifier. After removing the biasing circuit and using the common-emitter hybrid model for the transistor at low frequencies, an equivalent network of the amplifier is presented in Figure 14.7 with

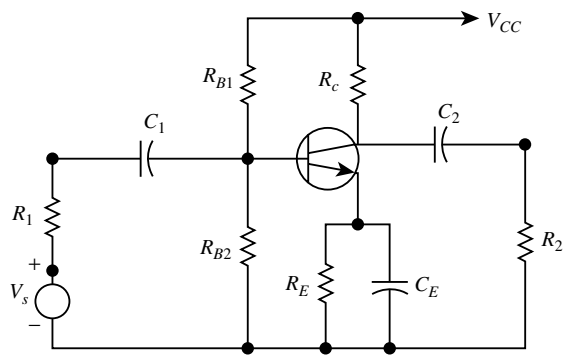


FIGURE 14.6 A common-emitter transistor feedback amplifier.

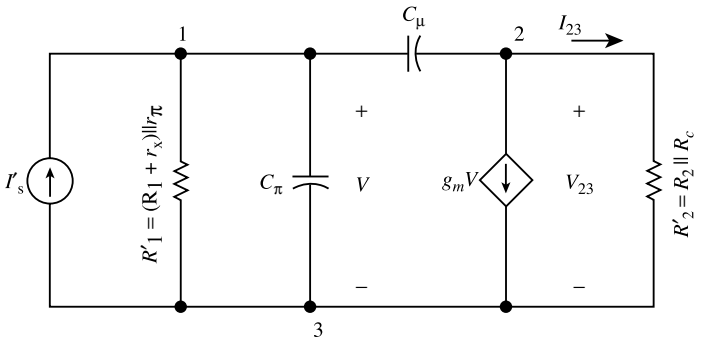


FIGURE 14.7 An equivalent network of the feedback amplifier of Figure 14.6.

$$I'_s = \frac{V_s}{R_1 + r_x} \tag{14.38a}$$

$$G'_1 = \frac{1}{R'_1} = \frac{1}{R_1 + r_x} + \frac{1}{r_\pi} \tag{14.38b}$$

$$G'_2 = \frac{1}{R'_2} = \frac{1}{R_2} + \frac{1}{R_c} \tag{14.38c}$$

The indefinite admittance matrix of the amplifier is:

$$\mathbf{Y} = \begin{bmatrix} G'_1 + sC_\pi + sC_\mu & -sC_\mu & -G'_1 - sC_\pi \\ g_m - sC_\mu & G'_2 + sC_\mu & -G'_2 - g_m \\ -G'_1 - sC_\pi - g_m & -G'_2 & G'_1 + G'_2 + sC_\pi + g_m \end{bmatrix} \tag{14.39}$$

Assume that the controlling parameter g_m is the element of interest. The return difference and the null return difference with respect to g_m in Figure 14.7 with I'_s as the input port and R'_2, as the output port, are:

$$F(g_m) = \frac{Y_{33}(g_m)}{Y_{33}(0)} = \frac{(G'_1 + sC_\pi)(G'_2 + sC_\mu) + sC_\mu(G'_2 + g_m)}{(G'_1 + sC_\pi)(G'_2 + sC_\mu) + sC_\mu G'_2} \tag{14.40}$$

$$\hat{F}(g_m) = \frac{Y_{12,33}(g_m)}{Y_{12,33}(0)} = \frac{sC_\mu - g_m}{sC_\mu} = 1 - \frac{g_m}{sC_\mu} \tag{14.41}$$

The current gain I_{23}/I'_s as defined in Figure 14.7, is computed as

$$w(g_m) = \frac{Y_{12,33}(g_m)}{R'_2 Y_{33}(g_m)} = \frac{sC_\mu - g_m}{R'_2\left[(G'_1 + sC_\pi)(G'_2 + sC_\mu) + sC_\mu(G'_2 + g_m)\right]} \tag{14.42}$$

Substituting these in (14.30) or (14.31) gives

$$\mathcal{S}(g_m) = -\frac{g_m(G'_1 + sC_\pi + sC_\mu)(G'_2 + sC_\mu)}{(sC_\mu - g_m)\left[(G'_1 + sC_\pi)(G'_2 + sC_\mu) + sC_\mu(G'_2 + g_m)\right]} \tag{14.43}$$

Finally, we compute the sensitivity for the driving-point impedance facing the current source I_s'. From (14.31), we obtain

$$\mathcal{S}\left(g_m\right)=\frac{1}{F\left(g_m\right)}\left[1-\frac{Z(0)}{Z\left(g_m\right)}\right]=-\frac{sC_\mu g_m}{\left(G_1'+sC_\pi\right)\left(G_2'+sC_\mu\right)+sC_\mu\left(G_2'+g_m\right)} \tag{14.44}$$

where

$$Z\left(g_m\right)=\frac{Y_{11,33}\left(g_m\right)}{Y_{33}\left(g_m\right)}=\frac{G_2'+sC_\mu}{\left(G_1'+sC_\pi\right)\left(G_2'+sC_\mu\right)+sC_\mu\left(G_2'+g_m\right)} \tag{14.45}$$

15

Measurement of Return Difference[1]

Wai-Kai Chen
University of Illinois, Chicago

The zeros of the network determinant are called the **natural frequencies**. Their locations in the complex-frequency plane are extremely important in that they determine the stability of the network. A network is said to be **stable** if all of its natural frequencies are restricted to the open left-half side (LHS) of the complex-frequency plane. If a network determinant is known, its roots can readily be computed explicitly with the aid of a computer if necessary, and the stability problem can then be settled directly. However, for a physical network there remains the difficulty of getting an accurate formulation of the network determinant itself, because every equivalent network is, to a greater or lesser extent, an idealization of the physical reality. As frequency is increased, parasitic effects of the physical elements must be taken into account. What is really needed is some kind of experimental verification that the network is stable and will remain so under certain prescribed conditions. The measurement of the return difference provides an elegant solution to this problem.

The return difference with respect to an element x in a feedback amplifier is defined by

$$F(x) = \frac{Y_{uv}(x)}{Y_{uv}(0)} \tag{15.1}$$

Because $Y_{uv}(x)$ denotes the nodal determinant, the zeros of the return difference are exactly the same as the zeros of the nodal determinant provided that there is no cancellation of common factors between $Y_{uv}(x)$ and $Y_{uv}(0)$. Therefore, if $Y_{uv}(0)$ is known to have no zeros in the closed right-half side (RHS) of the complex-frequency plane, which is usually the case in a single-loop feedback amplifier when x is set to zero, $F(x)$ gives precisely the same information about the stability of a feedback amplifier as does the nodal determinant itself. The difficulty inherent in the measurement of the return difference with respect to the controlling parameter of a controlled source is that, in a physical system, the controlling branch and the controlled source both form part of a single device such as a transistor, and cannot be physically separated. In the following, we present a scheme that does not require the physical decomposition of a device.

Let a device of interest be brought out as a two-port network connected to a general four-port network as shown in Figure 15.1. For our purposes, assume that this device is characterized by its y parameters, and represented by its y-parameter equivalent two-port network as indicated in Figure 15.2, in which

[1]References for this chapter can be found on page **16**-17.

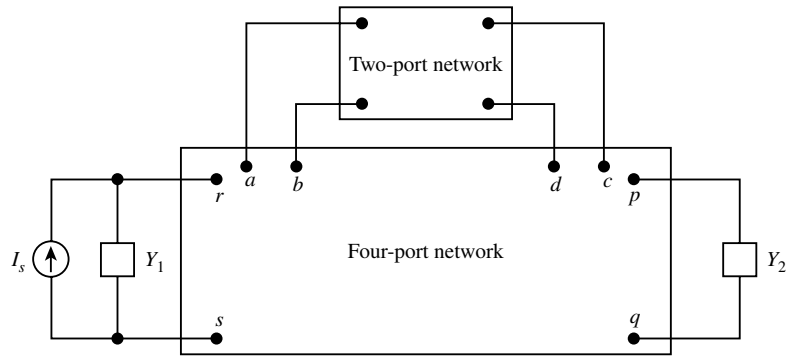


FIGURE 15.1 The general configuration of a feedback amplifier with a two-port device.

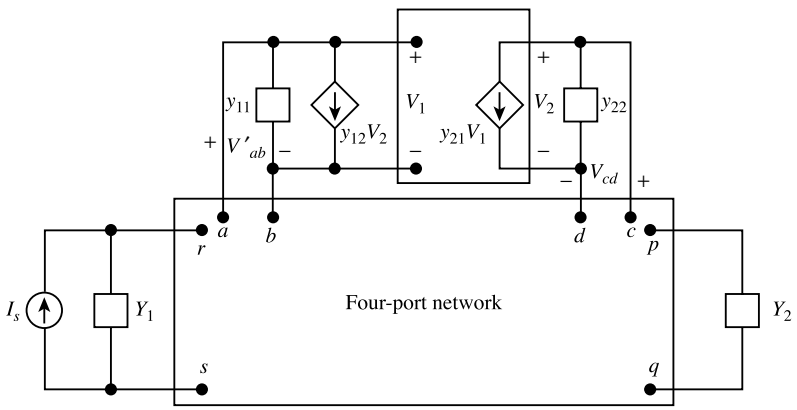


FIGURE 15.2 The representation of a two-port device in Figure 15.1 by its y parameters.

the parameter y_{21} controls signal transmission in the forward direction through the device, whereas y_{12} gives the reverse transmission, accounting for the internal feedback within the device. Our objective is to measure the return difference with respect to the forward short circuit transfer admittance y_{21}.

15.1 Blecher's Procedure [1]

Let the two-port device be a transistor operated in the common-emitter configuration with terminals a, $b = d$, and c representing, respectively, the base, emitter, and collector terminals. To simplify our notation, let $a = 1$, $b = d = 3$ and $c = 2$, as exhibited explicitly in Figure 15.3.

To measure $F(y_{21})$, we break the base terminal of the transistor and apply a 1-V excitation at its input as exhibited in Figure 15.3. To ensure that the controlled current source $y_{21}V_{13}$ drives a replica of what it sees during normal operation, we connect an active one-port network composed of a parallel combination of the admittance y_{11} and a controlled current source $y_{12}V_{23}$ at terminals 1 and 3. The returned voltage V_{13} is precisely the negative of the return ratio with respect to the element y_{21}. If, in the frequency band of interest, the externally applied feedback is large compared with the internal feedback of the transistor, the controlled source $y_{12}V_{23}$ can be ignored. If, however, we find that this internal feedback cannot be ignored, we can simulate it by using an additional transistor, connected as shown in Figure 15.4. This additional transistor must be matched as closely as possible to the one in question. The one-port admittance y_o denotes the admittance presented to the output port of the transistor under consideration as indicated in Figures 15.3 and 15.4. For a common-emitter state, it is perfectly reasonable to assume that $|y_o| \gg |y_{12}|$ and $|y_{11}| \gg |y_{12}|$. Under these assumptions, it is straightforward to show that the Norton equivalent network looking into the two-port network at terminals 1 and 3 of Figure 15.4 can be approximated by

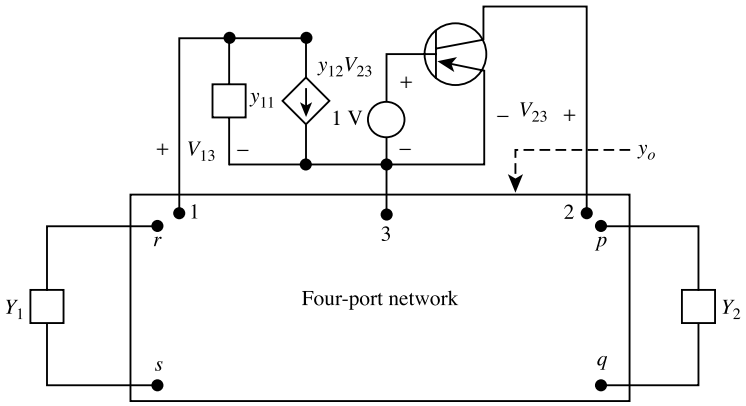


FIGURE 15.3 A physical interpretation of the return difference $F(y_{21})$ for a transistor operated in the common-emitter configuration and represented by its y parameters y_{ij}.

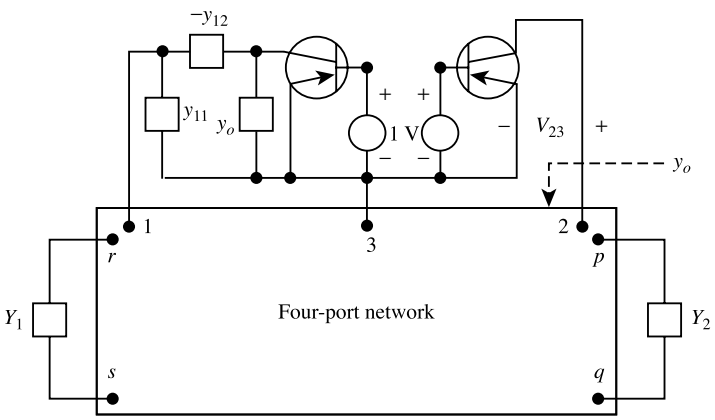


FIGURE 15.4 The measurement of return difference $F(y_{21})$ for a transistor operated in the common-emitter configuration and represented by its y parameters y_{ij}.

the parallel combination of y_{11} and $y_{12}V_{23}$, as indicated in Figure 15.3. In Figure 15.4, if the voltage sources have very low internal impedances, we can join together the two base terminals of the transistors and feed them both from a single voltage source of very low internal impedance. In this way, we avoid the need of using two separate sources. For the procedure to be feasible, we must demonstrate the admittances y_{11} and $-y_{12}$ can be realized as the input admittances of one-port RC networks.

Consider the hybrid-pi equivalent network of a common-emitter transistor of Figure 15.5, the short circuit admittance matrix of which is found to be

$$\mathbf{Y}_{sc} = \frac{1}{g_x + g_\pi + sC_\pi + sC_\mu} \begin{bmatrix} g_x\left(g_\pi + sC_\pi + sC_\mu\right) & -g_x sC_\mu \\ g_x\left(g_m - sC_\mu\right) & sC_\mu\left(g_x + g_\pi + sC_\pi + g_m\right) \end{bmatrix} \tag{15.2}$$

It is easy to confirm that the admittance y_{11} and $-y_{12}$ can be realized by the one-port networks of Figure 15.6.

15.2 Impedance Measurements

In this section, we show that the return difference can be evaluated by measuring two driving-point impedances at a convenient port in the feedback amplifier [8].

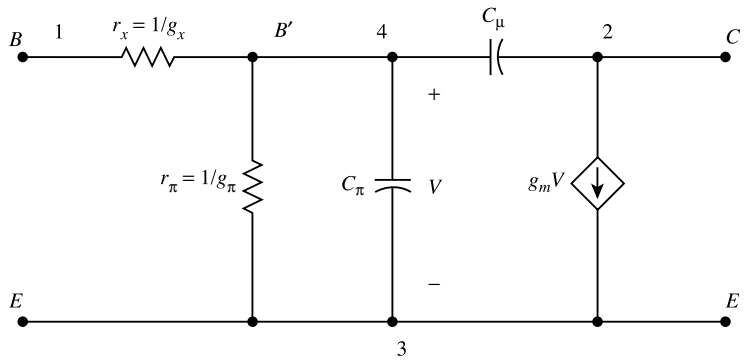


FIGURE 15.5 The hybrid-pi equivalent network of a common-emitter transistor.

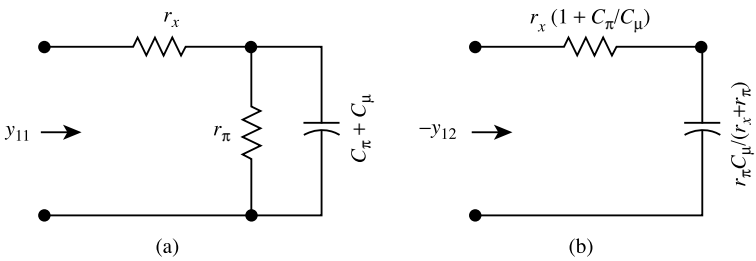


FIGURE 15.6 (a) The realization of y_{11} and (b) the realization of $-y_{12}$.

Refer again to the general feedback configuration of Figure 15.2. Suppose that we wish to evaluate the return difference with respect to the forward short circuit transfer admittance y_{21}. The controlling parameters y_{12} and y_{21} enter the indefinite-admittance matrix **Y** in the rectangular patterns as shown next:

$$\mathbf{Y}(x) = \begin{array}{c} \\ a \\ b \\ c \\ d \end{array} \begin{array}{c} \begin{array}{cccc} a & b & c & d \end{array} \\ \begin{bmatrix} & & y_{12} & -y_{12} \\ & & -y_{12} & y_{12} \\ y_{21} & -y_{21} & & \\ -y_{21} & y_{21} & & \end{bmatrix} \end{array} \tag{15.3}$$

To emphasize the importance of y_{12} and y_{21}, we again write $Y_{uv}(x)$ as $Y_{uv}(y_{12}, y_{21})$ and $z_{aa,bb}(x)$ as $z_{aa,bb}(y_{12}, y_{21})$. By appealing to formula (13.25), the impedance looking into terminals a and b of Figure 15.2 is:

$$z_{aa,bb}(y_{12}, y_{21}) = \frac{Y_{aa,bb}(y_{12}, y_{21})}{Y_{dd}(y_{12}, y_{21})} \tag{15.4}$$

The return difference with respect to y_{21} is given by

$$F(y_{21}) = \frac{Y_{dd}(y_{12}, y_{21})}{Y_{dd}(y_{12}, 0)} \tag{15.5}$$

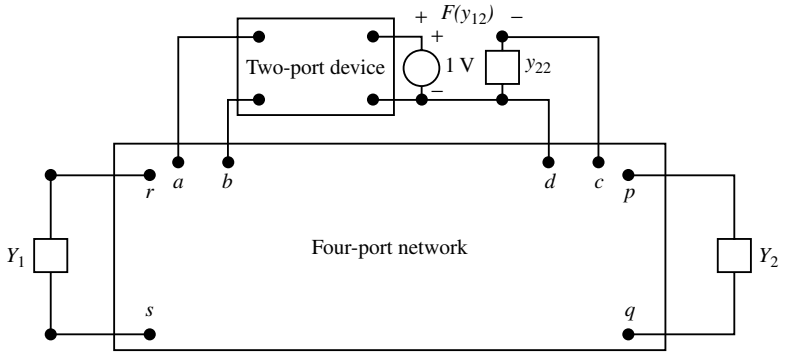


FIGURE 15.7 The measurement of the return difference $F(y_{12})$ with y_{21} set to zero.

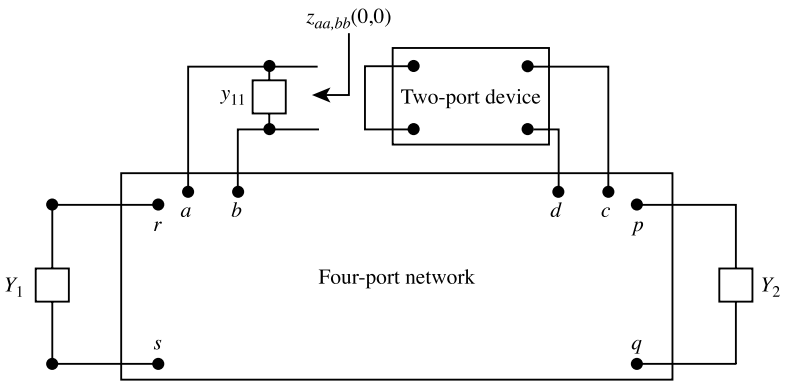


FIGURE 15.8 The measurement of the driving-point impedance $z_{aa,bb}(0, 0)$.

Combining these yields

$$F(y_{21})z_{aa,bb}(y_{12},y_{21}) = \frac{Y_{aa,bb}(y_{12},y_{21})}{Y_{dd}(y_{12},0)} = \frac{Y_{aa,bb}(0,0)}{Y_{dd}(y_{12},0)}$$
$$= \frac{Y_{aa,bb}(0,0)}{Y_{dd}(0,0)}\frac{Y_{dd}(0,0)}{Y_{dd}(y_{12},0)} = \frac{z_{aa,bb}(0,0)}{F(y_{12})\big|_{y_{21}=0}} \tag{15.6}$$

obtaining a relation

$$F(y_{12})\big|_{y_{21}=0} F(y_{21}) = \frac{z_{aa,bb}(0,0)}{z_{aa,bb}(y_{12},y_{21})} \tag{15.7}$$

among the return differences and the driving-point impedances. $F(y_{12})\big|_{y_{21}=0}$ is the return difference with respect to y_{12} when y_{21} is set to zero. This quantity can be measured by the arrangement of Figure 15.7. $z_{aa,bb}(y_{12}, y_{21})$ is the driving-point impedance looking into terminals a and b of the network of Figure 15.2. Finally, $z_{aa,bb}(0, 0)$ is the impedance to which $z_{aa,bb}(y_{12}, y_{21})$ reduces when the controlling parameters y_{12} and y_{21} are both set to zero. This impedance can be measured by the arrangement of Figure 15.8. Note that, in all three measurements, the independent current source I_s is removed.

Suppose that we wish to measure the return difference $F(y_{21})$ with respect to the forward transfer admittance y_{21} of a common-emitter transistor shown in Figure 15.2. Then, the return difference $F(y_{12})$

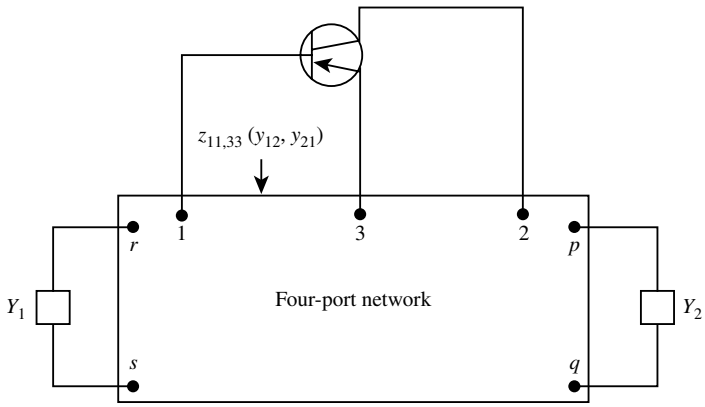


FIGURE 15.9 The measurement of the driving-point impedance $z_{11,33}(y_{12}, y_{21})$.

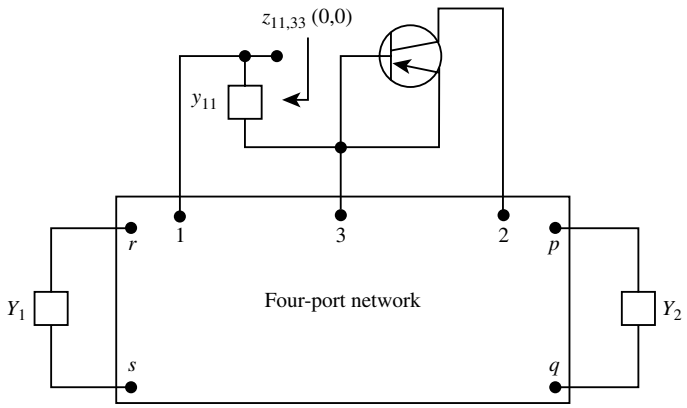


FIGURE 15.10 The measurement of the driving-point impedance $z_{11,33}(0, 0)$.

when y_{21} is set to zero, for all practical purposes, is indistinguishable from unity. Therefore, (15.7) reduces to the following simpler form:

$$F(y_{21}) \approx \frac{z_{11,33}(0,0)}{z_{11,33}(y_{12}, y_{21})} \tag{15.8}$$

showing that the return difference $F(y_{21})$ effectively equals the ratio of two functional values assumed by the driving-point impedance looking into terminals 1 and 3 of Figure 15.2 under the condition that the controlling parameters y_{12} and y_{21} are both set to zero and the condition that they assume their nominal values. These two impedances can be measured by the network arrangements of Figures 15.9 and 15.10.

16

Multiple-Loop Feedback Amplifiers

Wai-Kai Chen
University of Illinois, Chicago

So far, we have studied the single-loop feedback amplifiers. The concept of feedback was introduced in terms of return difference. We found that return difference is the difference between the unit applied signal and the returned signal. The returned signal has the same physical meaning as the loop transmission in the ideal feedback mode. It plays an important role in the study of amplifier stability, its sensitivity to the variations of the parameters, and the determination of its transfer and driving point impedances. The fact that return difference can be measured experimentally for many practical amplifiers indicates that we can include all the parasitic effects in the stability study, and that stability problem can be reduced to a Nyquist plot.

In this section, we study amplifiers that contain a multiplicity of inputs, outputs, and feedback loops. They are referred to as the **multiple-loop feedback amplifiers**. As might be expected, the notion of return difference with respect to an element is no longer applicable, because we are dealing with a group of elements. For this, we generalize the concept of return difference for a controlled source to the notion of return difference matrix for a multiplicity of controlled sources. For measurement situations, we introduce the null return difference matrix and discuss its physical significance. We demonstrate that the determinant of the overall transfer function matrix can be expressed explicity in terms of the determinants of the return difference and the null return difference matrices, thereby allowing us to generalize Blackman's formula for the input impedance.

16.1 Multiple-Loop Feedback Amplifier Theory

The general configuration of a multiple-input, multiple-output, and multiple-loop feedback amplifier is presented in Figure 16.1, in which the input, output, and feedback variables may be either currents or voltages. For the specific arrangement of Figure 16.1, the input and output variables are represented by an n-dimensional vector $\mathbf{u}$ and an m-dimensional vector $\mathbf{y}$ as

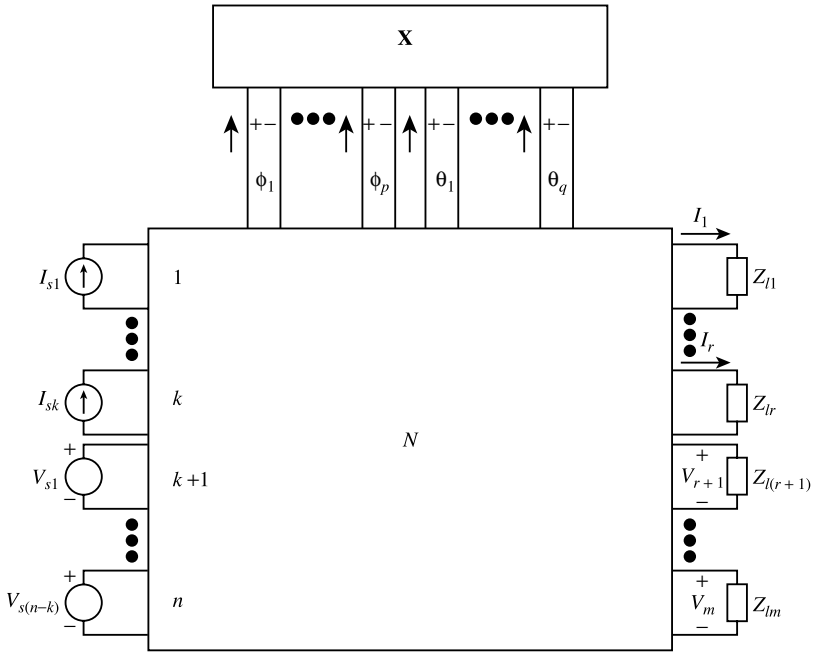


FIGURE 16.1 The general configuration of a multiple-input, multiple-output, and multiple-loop feedback amplifier.

$$\mathbf{u}(s)=\begin{bmatrix}u_1\\u_2\\\vdots\\u_k\\u_{k+1}\\u_{k+2}\\\vdots\\u_n\end{bmatrix}=\begin{bmatrix}I_{s1}\\I_{s2}\\\vdots\\I_{sk}\\V_{s1}\\V_{s2}\\\vdots\\V_{s(n-k)}\end{bmatrix},\quad \mathbf{y}(s)=\begin{bmatrix}y_1\\y_2\\\vdots\\y_r\\y_{r+1}\\y_{r+2}\\\vdots\\y_m\end{bmatrix}=\begin{bmatrix}I_1\\I_2\\\vdots\\I_r\\V_{r+1}\\V_{r+2}\\\vdots\\V_m\end{bmatrix} \tag{16.1}$$

respectively. The elements of interest can be represented by a rectangular matrix **X** of order $q \times p$ relating the controlled and controlling variables by the matrix equation

$$\mathbf{\Theta}=\begin{bmatrix}\theta_1\\\theta_2\\\vdots\\\theta_q\end{bmatrix}=\begin{bmatrix}x_{11}&x_{12}&\cdots&x_{1p}\\x_{21}&x_{22}&\cdots&x_{2p}\\\vdots&\vdots&\vdots&\vdots\\x_{q1}&x_{q2}&\cdots&x_{qp}\end{bmatrix}\begin{bmatrix}\phi_1\\\phi_2\\\vdots\\\phi_p\end{bmatrix}=\mathbf{X\Phi} \tag{16.2}$$

where the p-dimensional vector $\mathbf{\Phi}$ is called the **controlling vector**, and the q-dimensional vector $\mathbf{\Theta}$ is the **controlled vector**. The controlled variables θ_k and the controlling variables Φ_k can either be currents or voltages. The matrix **X** can represent either a transfer-function matrix or a driving-point function matrix.

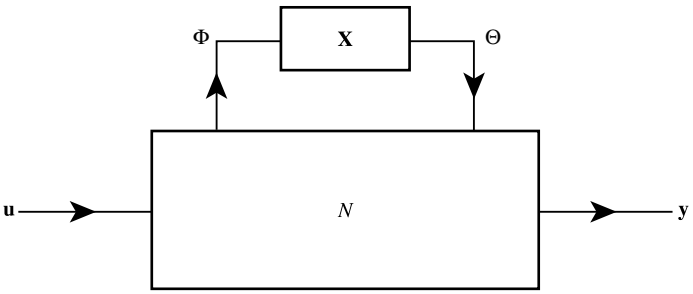


FIGURE 16.2 The block diagram of the general feedback configuration of Figure 16.1.

If **X** represents a driving-point function matrix, the vectors **Θ** and **Φ** are of the same dimension ($q = p$) and their components are the currents and voltages of a p-port network.

The general configuration of Figure 16.1 can be represented equivalently by the block diagram of Figure 16.2 in which N is a $(p + q + m + n)$-port network and the elements of interest are exhibited explicitly by the block **X**. For the $(p + q + m + n)$-port network N, the vectors **u** and are **Θ** are its inputs, and the vectors **Φ** and **y** its outputs. Since N is linear, the input and output vectors are related by the matrix equations

$$\mathbf{\Phi} = \mathbf{A\Theta} + \mathbf{Bu} \tag{16.3a}$$

$$\mathbf{y} = \mathbf{C\Theta} + \mathbf{Du} \tag{16.3b}$$

where **A**, **B**, **C**, and **D** are transfer-function matrices of orders $p \times q$, $p \times n$, $m \times q$, and $m \times n$, respectively. The vectors **Θ** and **Φ** are not independent and are related by

$$\mathbf{\Theta} = \mathbf{X\Phi} \tag{16.3c}$$

The relationships among the above three linear matrix equations can also be represented by a matrix signal-flow graph as shown in Figure 16.3 know as the **fundamental matrix feedback-flow graph**. The overall closed-loop **transfer-function matrix** of the multiple-loop feedback amplifier is defined by the equation

$$\mathbf{y} = \mathbf{W}(\mathbf{X})\mathbf{u} \tag{16.4}$$

where $\mathbf{W}(\mathbf{X})$ is of order $m \times n$. As before, to emphasize the importance of **X**, the matrix **W** is written as $\mathbf{W}(\mathbf{X})$ for the present discussion, even though it is also a function of the complex-frequency variable s. Combining the previous matrix equations, the transfer-function matrix is:

$$\mathbf{W}(\mathbf{X}) = \mathbf{D} + \mathbf{CX}\left(\mathbf{1}_p - \mathbf{AX}\right)^{-1}\mathbf{B} \tag{16.5a}$$

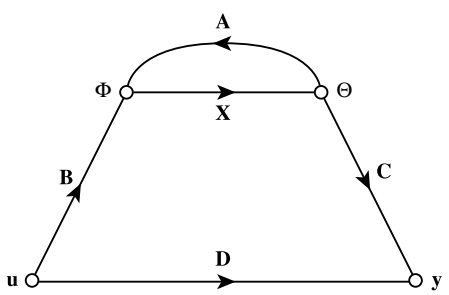


FIGURE 16.3 The fundamental matrix feedback-flow graph.

or

$$\mathbf{W}(\mathbf{X}) = \mathbf{D} + \mathbf{C}\left(\mathbf{1}_q - \mathbf{XA}\right)^{-1}\mathbf{XB} \tag{16.5b}$$

where $\mathbf{1}_p$ denotes the identity matrix of order p. Clearly, we have

$$\mathbf{W}(\mathbf{0}) = \mathbf{D} \tag{16.6}$$

In particular, when $\mathbf{X}$ is square and nonsingular, (16.5) can be written as

$$\mathbf{W}(\mathbf{X}) = \mathbf{D} + \mathbf{C}\left(\mathbf{X}^{-1} - \mathbf{A}\right)^{-1}\mathbf{B} \tag{16.7}$$

Example 3. Consider the voltage-series feedback amplifier of Figure 13.9. An equivalent network is shown in Figure 16.4 in which we have assumed that the two transistors are identical with $h_{ie} = 1.1$ kΩ, $h_{fe} = 50$, $h_{re} = h_{oe} = 0$. Let the controlling parameters of the two controlled sources be the elements of interest. Then we have

$$\mathbf{\Theta} = \begin{bmatrix} I_a \\ I_b \end{bmatrix} = 10^{-4}\begin{bmatrix} 455 & 0 \\ 0 & 455 \end{bmatrix}\begin{bmatrix} V_{13} \\ V_{45} \end{bmatrix} = \mathbf{X\Phi} \tag{16.8}$$

Assume that the output voltage V_{25} and input current I_{51} are the output variables. Then the seven-port network N defined by the variables V_{13}, V_{45}, V_{25}, I_{51}, I_a, I_b, and V_s can be characterized by the matrix equations

$$\begin{aligned} \mathbf{\Phi} &= \begin{bmatrix} V_{13} \\ V_{45} \end{bmatrix} = \begin{bmatrix} -90.782 & 45.391 \\ -942.507 & 0 \end{bmatrix}\begin{bmatrix} I_a \\ I_b \end{bmatrix} + \begin{bmatrix} 0.91748 \\ 0 \end{bmatrix}\begin{bmatrix} V_s \end{bmatrix} \\ &= \mathbf{A\Theta} + \mathbf{Bu} \end{aligned} \tag{16.9a}$$

$$\begin{aligned} \mathbf{y} &= \begin{bmatrix} V_{25} \\ I_{51} \end{bmatrix} = \begin{bmatrix} 45.391 & -2372.32 \\ -0.08252 & 0.04126 \end{bmatrix}\begin{bmatrix} I_a \\ I_b \end{bmatrix} + \begin{bmatrix} 0.041260 \\ 0.000862 \end{bmatrix}\begin{bmatrix} V_s \end{bmatrix} \\ &= \mathbf{C\Theta} + \mathbf{Du} \end{aligned} \tag{16.9b}$$

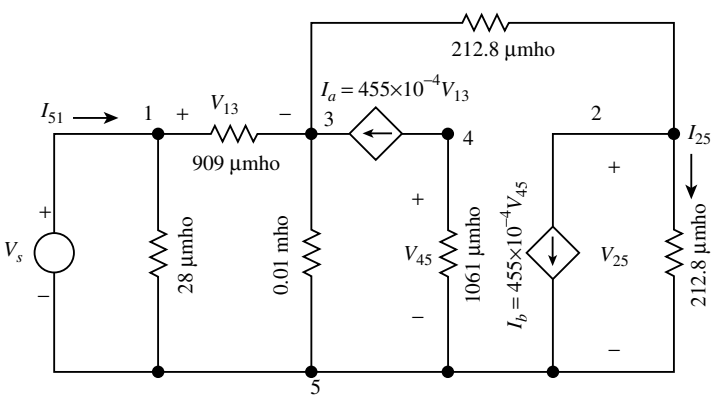


FIGURE 16.4 An equivalent network of the voltage-series feedback amplifier of Figure 13.9.

According to (16.4), the transfer-function matrix of the amplifier is defined by the matrix equation

$$\mathbf{y} = \begin{bmatrix} V_{25} \\ I_{51} \end{bmatrix} = \begin{bmatrix} w_{11} \\ w_{21} \end{bmatrix} \begin{bmatrix} V_s \end{bmatrix} = \mathbf{W}(\mathbf{X})\mathbf{u} \tag{16.10}$$

Because **X** is square and nonsingular, we can use (16.7) to calculate **W**(**X**):

$$\mathbf{W}(\mathbf{X}) = \mathbf{D} + \mathbf{C}\left(\mathbf{X}^{-1} - \mathbf{A}\right)^{-1}\mathbf{B} = \begin{bmatrix} 45.387 \\ 0.369 \times 10^{-4} \end{bmatrix} = \begin{bmatrix} w_{11} \\ w_{21} \end{bmatrix} \tag{16.11}$$

where

$$\left(\mathbf{X}^{-1} - \mathbf{A}\right)^{-1} = 10^{-4}\begin{bmatrix} 4.856 & 10.029 \\ -208.245 & 24.914 \end{bmatrix} \tag{16.12}$$

obtaining the closed-loop voltage gain w_{11} and input impedance Z_{in} facing the voltage source V_s as

$$w_{11} = \frac{V_{25}}{V_s} = 45.387, \quad Z_{\text{in}} = \frac{V_s}{I_{51}} = \frac{1}{w_{21}} = 27.1\ \mathbf{k\Omega} \tag{16.13}$$

16.2 The Return Different Matrix

In this section, we extend the concept of return difference with respect to an element to the notion of return difference matrix with respect to a group of elements.

In the fundamental matrix feedback-flow graph of Figure 16.3, suppose that we break the input of the branch with transmittance **X**, set the input excitation vector **u** to zero, and apply a signal p-vector **g** to the right of the breaking mark, as depicted in Figure 16.5. Then the returned signal p-vector **h** to the left of the breaking mark is found to be

$$\mathbf{h} = \mathbf{AXg} \tag{16.14}$$

The square matrix **AX** is called the **loop-transmission matrix** and its negative is referred to as the **return ratio matrix** denoted by

$$\mathbf{T}(\mathbf{X}) = -\mathbf{AX} \tag{16.15}$$

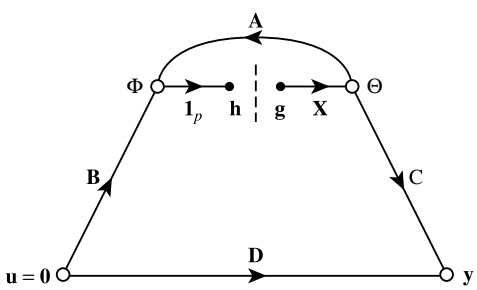


FIGURE 16.5 The physical interpretation of the loop-transmission matrix.

The difference between the applied signal vector **g** and the returned signal vector **h** is given by

$$\mathbf{g} - \mathbf{h} = \left(\mathbf{1}_p - \mathbf{AX}\right)\mathbf{g} \tag{16.16}$$

The square matrix $\mathbf{1}_p - \mathbf{AX}$ relating the applied signal vector **g** to the difference of the applied signal vector **g** and the returned signal vector **h** is called the **return difference matrix** with respect to **X** and is denoted by

$$\mathbf{F}(\mathbf{X}) = \mathbf{1}_p - \mathbf{AX} \tag{16.17}$$

Combining this with (16.15) gives

$$\mathbf{F}(\mathbf{X}) = \mathbf{1}_p + \mathbf{T}(\mathbf{X}) \tag{16.18}$$

For the voltage-series feedback amplifier of Figure 16.4, let the controlling parameters of the two controlled current sources be the elements of interest. Then the return ratio matrix is found from (16.8) and (16.9a)

$$\begin{aligned}\mathbf{T}(\mathbf{X}) = -\mathbf{AX} &= -\begin{bmatrix} -90.782 & 45.391 \\ -942.507 & 0 \end{bmatrix}\begin{bmatrix} 455\times10^{-4} & 0 \\ 0 & 455\times10^{-4} \end{bmatrix} \\ &= \begin{bmatrix} 4.131 & -2.065 \\ 42.884 & 0 \end{bmatrix}\end{aligned} \tag{16.19}$$

obtaining the return difference matrix as

$$\mathbf{F}(\mathbf{X}) = \mathbf{1}_2 + \mathbf{T}(\mathbf{X}) = \begin{bmatrix} 5.131 & -2.065 \\ 42.884 & 1 \end{bmatrix} \tag{16.20}$$

16.3 The Null Return Difference Matrix

A direct extension of the null return difference for the single-loop feedback amplifier is the null return difference matrix for the multiple-loop feedback networks.

Refer again to the fundamental matrix feedback-flow graph of Figure 16.3. As before, we break the branch with transmittance **X** and apply a signal p-vector **g** to the right of the breaking mark, as illustrated in Figure 16.6. We then adjust the input excitation n-vector **u** so that the total output m-vector **y** resulting from the inputs **g** and **u** is zero. From Figure 16.6, the desired input excitation **u** is found:

$$\mathbf{Du} + \mathbf{CXg} = \mathbf{0} \tag{16.21}$$

or

$$\mathbf{u} = -\mathbf{D}^{-1}\mathbf{CXg} \tag{16.22}$$

provided that the matrix **D** is square and nonsingular. This requires that the output **y** be of the same dimension as the input **u** or $m = n$. Physically, this requirement is reasonable because the effects at the output caused by **g** can be neutralized by a unique input excitation **u** only when **u** and **y** are of the same dimension. With these inputs **u** and **g**, the returned signal **h** to the left of the breaking mark in Figure 16.6 is computed as

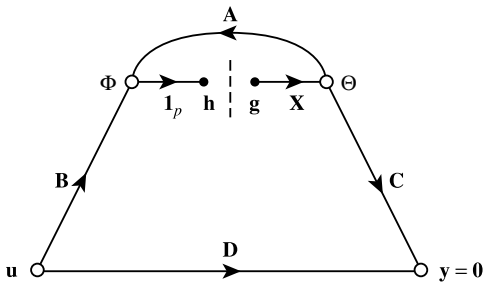


FIGURE 16.6 The physical interpretation of the null return difference matrix.

$$\mathbf{h} = \mathbf{Bu} + \mathbf{AXg} = \left(-\mathbf{BD}^{-1}\mathbf{CX} + \mathbf{AX}\right)\mathbf{g} \tag{16.23}$$

obtaining

$$\mathbf{g} - \mathbf{h} = \left(\mathbf{1}_p - \mathbf{AX} + \mathbf{BD}^{-1}\mathbf{CX}\right)\mathbf{g} \tag{16.24}$$

The square matrix

$$\hat{\mathbf{F}}(\mathbf{X}) = \mathbf{1}_p + \hat{\mathbf{T}}(\mathbf{X}) = \mathbf{1}_p - \mathbf{AX} + \mathbf{BD}^{-1}\mathbf{CX} = \mathbf{1}_p - \hat{\mathbf{A}}\mathbf{X} \tag{16.25}$$

relating the input signal vector **g** to the difference of the input signal vector **g**, and the returned signal vector **h** is called the **null return difference matrix** with respect to **X**, where

$$\hat{\mathbf{T}}(\mathbf{X}) = -\mathbf{AX} + \mathbf{BD}^{-1}\mathbf{CX} = -\hat{\mathbf{A}}\mathbf{X} \tag{16.26a}$$

$$\hat{\mathbf{A}} = \mathbf{A} - \mathbf{BD}^{-1}\mathbf{C} \tag{16.26b}$$

The square matrix $\hat{\mathbf{T}}(\mathbf{X})$ is known as the *null return ratio matrix.*

Example 4. Consider again the voltage-series feedback amplifier of Figure 13.9, an equivalent network of which is illustrated in Figure 16.4. Assume that the voltage V_{25} is the output variable. Then from (16.9)

$$\mathbf{\Phi} = \begin{bmatrix} V_{13} \\ V_{45} \end{bmatrix} = \begin{bmatrix} -90.782 & 45.391 \\ -942.507 & 0 \end{bmatrix} \begin{bmatrix} I_a \\ I_b \end{bmatrix} + \begin{bmatrix} 0.91748 \\ 0 \end{bmatrix} \begin{bmatrix} V_s \end{bmatrix} \tag{16.27a}$$

$$= \mathbf{A\Theta} + \mathbf{B}u$$

$$y = \begin{bmatrix} V_{25} \end{bmatrix} = \begin{bmatrix} 45.391 & -2372.32 \end{bmatrix} \begin{bmatrix} I_a \\ I_b \end{bmatrix} + \begin{bmatrix} 0.04126 \end{bmatrix} \begin{bmatrix} V_s \end{bmatrix} \tag{16.27b}$$

$$= \mathbf{C\Theta} + \mathbf{D}u$$

Substituting the coefficient matrices in (16.26b), we obtain

$$\hat{\mathbf{A}} = \mathbf{A} - \mathbf{BD}^{-1}\mathbf{C} = \begin{bmatrix} -1100.12 & 52{,}797.6 \\ -942.507 & 0 \end{bmatrix} \tag{16.28}$$

giving the null return difference matrix with respect to **X** as

$$\hat{\mathbf{F}}(\mathbf{X}) = \mathbf{1}_2 - \hat{\mathbf{A}}\mathbf{X} = \begin{bmatrix} 51.055 & -2402.29 \\ 42.884 & 1 \end{bmatrix} \tag{16.29}$$

Suppose that the input current I_{51} is chosen as the output variable. Then, from (16.9b) we have

$$y = \begin{bmatrix} I_{51} \end{bmatrix} = \begin{bmatrix} -0.08252 & 0.04126 \end{bmatrix} \begin{bmatrix} I_a \\ I_b \end{bmatrix} + \begin{bmatrix} 0.000862 \end{bmatrix} \begin{bmatrix} V_s \end{bmatrix} = \mathbf{C}\boldsymbol{\Theta} + Du \tag{16.30}$$

The corresponding null return difference matrix becomes

$$\hat{\mathbf{F}}(\mathbf{X}) = \mathbf{1}_2 - \hat{\mathbf{A}}\mathbf{X} = \begin{bmatrix} 1.13426 & -0.06713 \\ 42.8841 & 1 \end{bmatrix} \tag{16.31}$$

where

$$\hat{\mathbf{A}} = \begin{bmatrix} -2.95085 & 1.47543 \\ -942.507 & 0 \end{bmatrix} \tag{16.32}$$

16.4 The Transfer-Function Matrix and Feedback

In this section, we show the effect of feedback on the transfer-function matrix $\mathbf{W}(\mathbf{X})$. Specifically, we express $\det \mathbf{W}(\mathbf{X})$ in terms of the $\det \mathbf{X}(\mathbf{0})$ and the determinants of the return difference and null return difference matrices, thereby generalizing Blackman's impedance formula for a single input to a multiplicity of inputs.

Before we proceed to develop the desired relation, we state the following determinant identity for two arbitrary matrices **M** and **N** of order $m \times n$ and $n \times m$:

$$\det(\mathbf{1}_m + \mathbf{MN}) = \det(\mathbf{1}_n + \mathbf{NM}) \tag{16.33}$$

a proof of which may be found in [5, 6]. Using this, we next establish the following generalization of Blackman's formula for input impedance.

Theorem 1. *In a multiple-loop feedback amplifier, if* $\mathbf{W}(\mathbf{0}) = \mathbf{D}$ *is nonsingular, then the determinant of the transfer-function matrix* $\mathbf{W}(\mathbf{X})$ *is related to the determinants of the return difference matrix* $\mathbf{F}(\mathbf{X})$ *and the null return difference matrix* $\hat{\mathbf{F}}(\mathbf{X})$ *by*

$$\det \mathbf{W}(\mathbf{X}) = \det \mathbf{W}(\mathbf{0}) \frac{\det \hat{\mathbf{F}}(\mathbf{X})}{\det \mathbf{F}(\mathbf{X})} \tag{16.34}$$

PROOF: From (16.5a), we obtain

$$\mathbf{W}(\mathbf{X}) = \mathbf{D}\left[\mathbf{1}_n + \mathbf{D}^{-1}\mathbf{CX}(\mathbf{1}_p - \mathbf{AX})^{-1}\mathbf{B}\right] \tag{16.35}$$

yielding

$$\begin{aligned}\det \mathbf{W}(\mathbf{X}) &= \left[\det \mathbf{W}(\mathbf{0})\right] \det\left[\mathbf{1}_n + \mathbf{D}^{-1}\mathbf{C}\mathbf{X}\left(\mathbf{1}_p - \mathbf{A}\mathbf{X}\right)^{-1}\mathbf{B}\right] \\ &= \left[\det \mathbf{W}(\mathbf{0})\right] \det\left[\mathbf{1}_p + \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\mathbf{X}\left(\mathbf{1}_p - \mathbf{A}\mathbf{X}\right)^{-1}\right] \\ &= \left[\det \mathbf{W}(\mathbf{0})\right] \det\left[\mathbf{1}_p - \mathbf{A}\mathbf{X} + \mathbf{B}\mathbf{D}^{-1}\mathbf{C}\mathbf{X}\right]\left(\mathbf{1}_p - \mathbf{A}\mathbf{X}\right)^{-1} \\ &= \frac{\det \mathbf{W}(\mathbf{0}) \det \hat{\mathbf{F}}(\mathbf{X})}{\det \mathbf{F}(\mathbf{X})}\end{aligned} \tag{16.36}$$

The second line follows directly from (16.33). This completes the proof of the theorem.

As indicated in (14.4), the input impedance $Z(x)$ looking into a terminal pair can be conveniently expressed as

$$Z(x) = Z(0)\frac{F(\text{input short-ciruited})}{F(\text{input open-circuited})} \tag{16.37}$$

A similar expression can be derived from (16.34) if $\mathbf{W}(\mathbf{X})$ denotes the impedance matrix of an n-port network of Figure 16.1. In this case, $\mathbf{F}(\mathbf{X})$ is the return difference matrix with respect to $\mathbf{X}$ for the situation when the n ports where the impedance matrix are defined are left open without any sources, and we write $\mathbf{F}(\mathbf{X}) = \mathbf{F}(\text{input open-circuited})$. Likewise, $\hat{\mathbf{F}}(\mathbf{X})$ is the return difference matrix with respect to $\mathbf{X}$ for the input port-current vector $\mathbf{I}_s$ and the output port-voltage vector $\mathbf{V}$ under the condition that $\mathbf{I}_s$ is adjusted so that the port-voltage vector $\mathbf{V}$ is identically zero. In other words, $\hat{\mathbf{F}}(\mathbf{X})$ is the return difference matrix for the situation when the n ports, where the impedance matrix is defined, are short-circuited, and we write $\hat{\mathbf{F}}(\mathbf{X}) = \mathbf{F}(\text{input short-circuited})$. Consequently, the determinant of the impedance matrix $\mathbf{Z}(\mathbf{X})$ of an n-port network can be expressed from (16.34) as

$$\det \mathbf{Z}(\mathbf{X}) = \det \mathbf{Z}(\mathbf{0})\frac{\det \mathbf{F}(\text{input short-circuited})}{\det \mathbf{F}(\text{input open-circuited})} \tag{16.38}$$

Example 5. Refer again to the voltage-series feedback amplifier of Figure 13.9, an equivalent network of which is illustrated in Figure 16.4. As computed in (16.20), the return difference matrix with respect to the two controlling parameters is given by

$$\mathbf{F}(\mathbf{X}) = \mathbf{1}_2 + \mathbf{T}(\mathbf{X}) = \begin{bmatrix} 5.131 & -2.065 \\ 42.884 & 1 \end{bmatrix} \tag{16.39}$$

the determinant of which is:

$$\det \mathbf{F}(\mathbf{X}) = 93.68646 \tag{16.40}$$

If V_{25} of Figure 16.4 is chosen as the output and V_s as the input, the null return difference matrix is, from (16.29),

$$\hat{\mathbf{F}}(\mathbf{X}) = \mathbf{1}_2 - \hat{\mathbf{A}}\mathbf{X} = \begin{bmatrix} 51.055 & -2402.29 \\ 42.884 & 1 \end{bmatrix} \tag{16.41}$$

the determinant of which is:

$$\det \hat{\mathbf{F}}(\mathbf{X}) = 103{,}071 \tag{16.42}$$

By appealing to (16.34), the feedback amplifier voltage gain V_{25}/V_s can be written as

$$w(\mathbf{X}) = \frac{V_{25}}{V_s} = w(\mathbf{0})\frac{\det \hat{\mathbf{F}}(\mathbf{X})}{\det \mathbf{F}(\mathbf{X})} = 0.04126\frac{103{,}071}{93.68646} = 45.39 \tag{16.43}$$

confirming (13.44), where $w(\mathbf{0}) = 0.04126$, as given in (16.27b).

Suppose, instead, that the input current I_{51} is chosen as the output and V_s as the input. Then, from (16.31), the null return difference matrix becomes

$$\hat{\mathbf{F}}(\mathbf{X}) = \mathbf{1}_2 - \hat{\mathbf{A}}(\mathbf{X}) = \begin{bmatrix} 1.13426 & -0.06713 \\ 42.8841 & 1 \end{bmatrix} \tag{16.44}$$

the determinant of which is:

$$\det \hat{\mathbf{F}}(\mathbf{X}) = 4.01307 \tag{16.45}$$

By applying (16.34), the amplifier input admittance is obtained as

$$\begin{aligned} w(\mathbf{X}) &= \frac{I_{51}}{V_s} = w(\mathbf{0})\frac{\det \hat{\mathbf{F}}(\mathbf{X})}{\det \mathbf{F}(\mathbf{X})} \\ &= 8.62 \times 10^{-4}\,\frac{4.01307}{93.68646} = 36.92\ \mu\text{mho} \end{aligned} \tag{16.46}$$

or 27.1 kΩ, confirming (16.13), where $w(\mathbf{0}) = 862$ μmho is found from (16.30).

Another useful application of the generalized Blackman's formula (16.38) is that it provides the basis of a procedure for the indirect measurement of return difference. Refer to the general feedback network of Figure 16.2. Suppose that we wish to measure the return difference $F(y_{21})$ with respect to the forward short circuit transfer admittance y_{21} of a two-port device characterized by its y parameters y_{ij}. Choose the two controlling parameters y_{21} and y_{12} to be the elements of interest. Then, from Figure 15.2 we obtain

$$\mathbf{\Theta} = \begin{bmatrix} I_a \\ I_b \end{bmatrix} = \begin{bmatrix} y_{21} & 0 \\ 0 & y_{12} \end{bmatrix}\begin{bmatrix} V_1 \\ V_2 \end{bmatrix} = \mathbf{X\Phi} \tag{16.47}$$

where I_a and I_b are the currents of the voltage-controlled current sources. By appealing to (16.38), the impedance looking into terminals a and b of Figure 15.2 can be written as

$$z_{aa,bb}(y_{12}, y_{21}) = z_{aa,bb}(0,0)\frac{\det \mathbf{F}(\text{input short-circuited})}{\det \mathbf{F}(\text{input open-circuited})} \tag{16.48}$$

When the input terminals a and b are open-circuited, the resulting return difference matrix is exactly the same as that found under normal operating conditions, and we have

$$\mathbf{F}(\text{input open-circuited}) = \mathbf{F}(\mathbf{X}) = \begin{bmatrix} F_{11} & F_{12} \\ F_{21} & F_{22} \end{bmatrix} \tag{16.49}$$

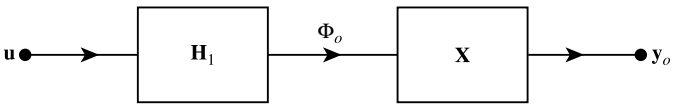


FIGURE 16.7 The block diagram of a multivariable open-loop control system.

Because

$$\mathbf{F}(\mathbf{X}) = \mathbf{1}_2 - \mathbf{AX} \tag{16.50}$$

the elements F_{11} and F_{21} are calculated with $y_{12} = 0$, whereas F_{12} and F_{22} are evaluated with $y_{21} = 0$. When the input terminals a and b are short circuited, the feedback loop is interrupted and only the second row and first column element of the matrix **A** is nonzero, and we obtain

$$\det \mathbf{F}(\text{input short-circuited}) = 1 \tag{16.51}$$

Because **X** is diagonal, the return difference function $F(y_{21})$ can be expressed in terms of det **F**(**X**) and the cofactor of the first row and first column element of **F**(**X**):

$$F(y_{21}) = \frac{\det \mathbf{F}(\mathbf{X})}{F_{22}} \tag{16.52}$$

Substituting these in (16.48) yields

$$F(y_{12})\Big|_{y_{21}=0} F(y_{21}) = \frac{z_{aa,bb}(0,0)}{z_{aa,bb}(y_{12,}y_{21})} \tag{16.53}$$

where

$$F_{22} = 1 - a_{22}y_{12}\Big|_{y_{21}=0} = F(y_{12})\Big|_{y_{21}=0} \tag{16.54}$$

and a_{22} is the second row and second column element of **A**. Formula (16.53) was derived earlier in (15.7) using the network arrangements of Figures 15.7 and 15.8 to measure the elements $F(y_{12})|_{y_{21}=0}$ and $z_{aa,bb}(0,0)$, respectively.

16.5 The Sensitivity Matrix

We have studied the sensitivity of a transfer function with respect to the change of a particular element in the network. In a multiple-loop feedback network, we are usually interested in the sensitivity of a transfer function with respect to the variation of a set of elements in the network. This set may include either elements that are inherently sensitive to variation or elements where the effect on the overall amplifier performance is of paramount importance to the designers. For this, we introduce a sensitivity matrix and develop formulas for computing multiparameter sensitivity function for a multiple-loop feedback amplifier [7].

Figure 16.7 is the block diagram of a multivariable open-loop control system with n inputs and m outputs, whereas Figure 16.8 is the general feedback structure. If all feedback signals are obtainable from the output and if the controllers are linear, no loss of generality occurs by assuming the controller to be of the form given in Figure 16.9.

Denote the set of Laplace-transformed input signals by the n-vector **u**, the set of inputs to the network **X** in the open-loop configuration of Figure 16.7 by the p-vector $\mathbf{\Phi}_o$, and the set of outputs of the network

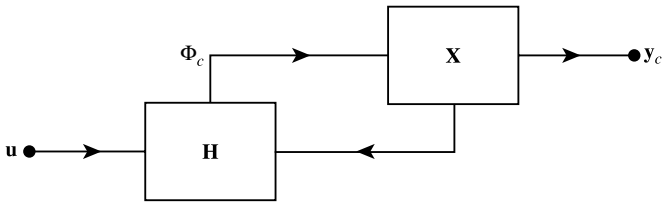


FIGURE 16.8 The general feedback structure.

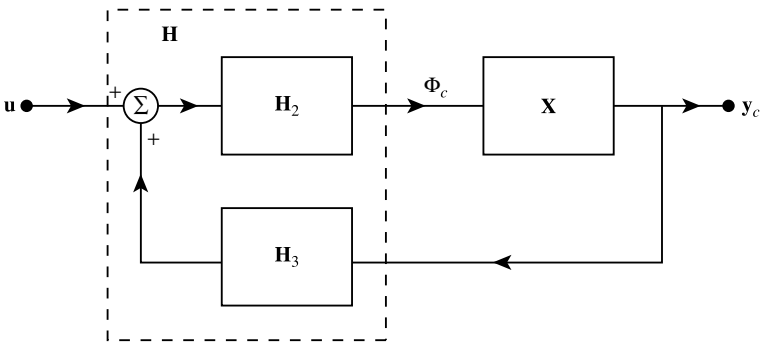


FIGURE 16.9 The general feedback configuration.

X of Figure 16.7 by the m-vector $\mathbf{y}_o$. Let the corresponding signals for the closed-loop configuration of Figure 16.9 be denoted by the n-vector **u**, the p-vector $\mathbf{\Phi}_c$, and the m-vector $\mathbf{y}_c$, respectively. Then, from Figures 16.7 and 16.9, we obtain the following relations:

$$\mathbf{y}_o = \mathbf{X}\mathbf{\Phi}_o \tag{16.55a}$$

$$\mathbf{\Phi}_o = \mathbf{H}_1\mathbf{u} \tag{16.55b}$$

$$\mathbf{y}_c = \mathbf{X}\mathbf{\Phi}_c \tag{16.55c}$$

$$\mathbf{\Phi}_c = \mathbf{H}_2\left(\mathbf{u} + \mathbf{H}_3\mathbf{y}_c\right) \tag{16.55d}$$

where the transfer-function matrices **X**, $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ are of order $m \times p$, $p \times n$, $p \times n$ and $n \times m$, respectively. Combining (16.55c) and (16.55d) yields

$$\left(\mathbf{1}_m - \mathbf{X}\mathbf{H}_2\mathbf{H}_3\right)\mathbf{y}_c = \mathbf{X}\mathbf{H}_2\mathbf{u} \tag{16.56}$$

or

$$\mathbf{y}_c = \left(\mathbf{1}_m - \mathbf{X}\mathbf{H}_2\mathbf{H}_3\right)^{-1}\mathbf{X}\mathbf{H}_2\mathbf{u} \tag{16.57}$$

The closed-loop transfer function matrix $\mathbf{W}(\mathbf{X})$ that relates the input vector **u** to the output vector $\mathbf{y}_c$ is defined by the equation

$$\mathbf{y}_c = \mathbf{W}(\mathbf{X})\mathbf{u} \tag{16.58}$$

identifying from (16.57) the $m \times n$ matrix

$$\mathbf{W}(\mathbf{X}) = \left(\mathbf{1}_m - \mathbf{X}\mathbf{H}_2\mathbf{H}_3\right)^{-1}\mathbf{X}\mathbf{H}_2 \tag{16.59}$$

Now, suppose that **X** is perturbed from **X** to **X** + **δX**. The outputs of the open-loop and closed-loop systems of Figure 16.7 and 16.9 will no longer be the same as before. Distinguishing the new from the old variables by the superscript +, we have

$$\mathbf{y}_o^+ = \mathbf{X}^+\mathbf{\Phi}_o \tag{16.60a}$$

$$\mathbf{y}_c^+ = \mathbf{X}^+\mathbf{\Phi}_c^+ \tag{16.60b}$$

$$\mathbf{\Phi}_c^+ = \mathbf{H}_2\left(\mathbf{u} + \mathbf{H}_3\mathbf{y}_c^+\right) \tag{16.60c}$$

where $\mathbf{\Phi}_o$ remains the same.

We next proceed to compare the relative effects of the variations of **X** on the performance of the open-loop and the closed-loop systems. For a meaningful comparison, we assume that $\mathbf{H}_1$, $\mathbf{H}_2$, and $\mathbf{H}_3$ are such that when there is no variation of **X**, $\mathbf{y}_o = \mathbf{y}_c$. Define the error vectors resulting from perturbation of **X** as

$$\mathbf{E}_o = \mathbf{y}_o - \mathbf{y}_o^+ \tag{16.61a}$$

$$\mathbf{E}_c = \mathbf{y}_c - \mathbf{y}_c^+ \tag{16.61b}$$

A square matrix relating $\mathbf{E}_o$ to $\mathbf{E}_c$ is called the **sensitivity matrix** $\mathcal{S}(\mathbf{X})$ for the transfer function matrix $\mathbf{W}(\mathbf{X})$ with respect to the variations of **X**:

$$\mathbf{E}_c = \mathcal{S}(\mathbf{X})\mathbf{E}_o \tag{16.62}$$

In the following, we express the **sensitivity matrix** $\mathcal{S}(\mathbf{X})$ in terms of the system matrices **X**, $\mathbf{H}_2$, and $\mathbf{H}_3$.

The input and output relation similar to that given in (16.57) for the perturbed system can be written as

$$\mathbf{y}_c^+ = \left(\mathbf{1}_m - \mathbf{X}^+\mathbf{H}_2\mathbf{H}_3\right)^{-1}\mathbf{X}^+\mathbf{H}_2\mathbf{u} \tag{16.63}$$

Substituting (16.57) and (16.63) in (16.61b) gives

$$\begin{aligned}
\mathbf{E}_c = \mathbf{y}_c - \mathbf{y}_c^+ &= \left[\left(\mathbf{1}_m - \mathbf{X}\mathbf{H}_2\mathbf{H}_3\right)^{-1}\mathbf{X}\mathbf{H}_2 - \left(\mathbf{1}_m - \mathbf{X}^+\mathbf{H}_2\mathbf{H}_3\right)^{-1}\mathbf{X}^+\mathbf{H}_2\right]\mathbf{u} \\
&= \left(\mathbf{1}_m - \mathbf{X}^+\mathbf{H}_2\mathbf{H}_3\right)^{-1}\left\{\left[\mathbf{1}_m - \left(\mathbf{X} + \mathbf{\delta X}\right)\mathbf{H}_2\mathbf{H}_3\right]\left(\mathbf{1}_m - \mathbf{X}\mathbf{H}_2\mathbf{H}_3\right)^{-1}\mathbf{X}\mathbf{H}_2 - \left(\mathbf{X} + \mathbf{\delta X}\right)\mathbf{H}_2\right\}\mathbf{u} \\
&= \left(\mathbf{1}_m - \mathbf{X}^+\mathbf{H}_2\mathbf{H}_3\right)^{-1}\left[\mathbf{X}\mathbf{H}_2 - \mathbf{\delta X}\mathbf{H}_2\mathbf{H}_3\left(\mathbf{1}_m - \mathbf{X}\mathbf{H}_2\mathbf{H}_3\right)^{-1}\mathbf{X}\mathbf{H}_2 - \mathbf{X}\mathbf{H}_2 - \mathbf{\delta X}\mathbf{H}_2\right]\mathbf{u} \\
&= -\left(\mathbf{1}_m - \mathbf{X}^+\mathbf{H}_2\mathbf{H}_3\right)^{-1}\mathbf{\delta X}\mathbf{H}_2\left[\mathbf{1}_n + \mathbf{H}_3\mathbf{W}(\mathbf{X})\right]\mathbf{u}
\end{aligned} \tag{16.64}$$

From (16.55d) and (16.58), we obtain

$$\mathbf{\Phi}_c = \mathbf{H}_2\left[\mathbf{1}_n + \mathbf{H}_3\mathbf{W}(\mathbf{X})\right]\mathbf{u} \tag{16.65}$$

Because by assuming that $\mathbf{y}_o = \mathbf{y}_c$, we have

$$\mathbf{\Phi}_o = \mathbf{\Phi}_c = \mathbf{H}_2\left[\mathbf{1}_n + \mathbf{H}_3\mathbf{W}(\mathbf{X})\right]\mathbf{u} \tag{16.66}$$

yielding

$$\mathbf{E}_o = \mathbf{y}_o - \mathbf{y}_o^+ = \left(\mathbf{X} - \mathbf{X}^+\right)\boldsymbol{\Phi}_o = -\boldsymbol{\delta}\mathbf{X}\mathbf{H}_2\left[\mathbf{1}_n + \mathbf{H}_3\mathbf{W}(\mathbf{X})\right]\mathbf{u} \tag{16.67}$$

Combining (16.64) and (16.67) yields an expression relating the error vectors $\mathbf{E}_c$ and $\mathbf{E}_o$ of the closed-loop and open-loop systems by

$$\mathbf{E}_c = \left(\mathbf{1}_m - \mathbf{X}^+\mathbf{H}_2\mathbf{H}_3\right)^{-1}\mathbf{E}_o \tag{16.68}$$

obtaining the sensitivity matrix as

$$\mathcal{S}(\mathbf{X}) = \left(\mathbf{1}_m - \mathbf{X}^+\mathbf{H}_2\mathbf{H}_3\right)^{-1} \tag{16.69}$$

For small variations of $\mathbf{X}$, $\mathbf{X}^+$ is approximately equal to $\mathbf{X}$. Thus, in Figure 16.9, if the matrix triple product $\mathbf{X}\mathbf{H}_2\mathbf{H}_3$ is regarded as the *loop-transmission matrix* and $-\mathbf{X}\mathbf{H}_2\mathbf{H}_3$ as the *return ratio matrix*, then the difference between the unit matrix and the loop-transmission matrix,

$$\mathbf{1}_m - \mathbf{X}\mathbf{H}_2\mathbf{H}_3 \tag{16.70}$$

can be defined as the *return difference matrix*. Therefore, (16.69) is a direct extension of the sensitivity function defined for a single-input, single-output system and for a single parameter. Recall that in (14.33) we demonstrated that, using the ideal feedback model, the sensitivity function of the closed-loop transfer function with respect to the forward amplifier gain is equal to the reciprocal of its return difference with respect to the same parameter.

In particular, when $\mathbf{W}(\mathbf{X})$, $\boldsymbol{\delta}\mathbf{X}$, and $\mathbf{X}$ are square and nonsingular, from (16.55a), (16.55b), and (16.58), (16.61) can be rewritten as

$$\mathbf{E}_c = \mathbf{y}_c - \mathbf{y}_c^+ = \left[\mathbf{W}(\mathbf{X}) - \mathbf{W}^+(\mathbf{X})\right]\mathbf{u} = -\boldsymbol{\delta}\mathbf{W}(\mathbf{X})\mathbf{u} \tag{16.71a}$$

$$\mathbf{E}_o = \mathbf{y}_o - \mathbf{y}_o^+ = \left[\mathbf{X}\mathbf{H}_1 - \mathbf{X}^+\mathbf{H}_1\right]\mathbf{u} = -\boldsymbol{\delta}\mathbf{X}\mathbf{H}_1\mathbf{u} \tag{16.71b}$$

If $\mathbf{H}_1$ is nonsingular, $\mathbf{u}$ in (16.71b) can be solved for and substituted in (16.71a) to give

$$\mathbf{E}_c = \boldsymbol{\delta}\mathbf{W}(\mathbf{X})\mathbf{H}_1^{-1}(\boldsymbol{\delta}\mathbf{X})^{-1}\mathbf{E}_o \tag{16.72}$$

As before, for meaningful comparison, we require that $\mathbf{y}_o = \mathbf{y}_c$ or

$$\mathbf{X}\mathbf{H}_1 = \mathbf{W}(\mathbf{X}) \tag{16.73}$$

From (16.72), we obtain

$$\mathbf{E}_c = \boldsymbol{\delta}\mathbf{W}(\mathbf{X})\mathbf{W}^{-1}(\mathbf{X})\mathbf{X}(\boldsymbol{\delta}\mathbf{X})^{-1}\mathbf{E}_o \tag{16.74}$$

identifying that

$$\mathcal{S}(\mathbf{X}) = \boldsymbol{\delta}\mathbf{W}(\mathbf{X})\mathbf{W}^{-1}(\mathbf{X})\mathbf{X}(\boldsymbol{\delta}\mathbf{X})^{-1} \tag{16.75}$$

This result is to be compared with the scalar sensitivity function defined in (14.26), which can be put in the form

$$\mathcal{S}(x) = (\delta w)w^{-1}x(\delta x)^{-1} \tag{16.76}$$

16.6 Multiparameter Sensitivity

In this section, we derive formulas for the effect of change of **X** on a scalar transfer function $w(\mathbf{X})$.

Let x_k, $k = 1, 2, \ldots, pq$, be the elements of **X**. The multivariable Taylor series expansion of $w(\mathbf{X})$ with respect to x_k is given by

$$\delta w = \sum_{k=1}^{pq} \frac{\partial w}{\partial x_k} \delta x_k + \sum_{j=1}^{pq} \sum_{k=1}^{pq} \frac{\partial^2 w}{\partial x_j \partial x_k} \frac{\delta x_j \delta x_k}{2!} + \cdots \tag{16.77}$$

The first-order perturbation can then be written as

$$\delta w \approx \sum_{k=1}^{pq} \frac{\partial w}{\partial x_k} \delta x_k \tag{16.78}$$

Using (14.26), we obtain

$$\frac{\delta w}{w} \approx \sum_{k=1}^{pq} \mathcal{S}(x_k) \frac{\delta x_k}{x_k} \tag{16.79}$$

This expression gives the fractional change of the transfer function w in terms of the scalar sensitivity functions $\mathcal{S}(x_k)$.

Refer to the fundamental matrix feedback-flow graph of Figure 16.3. If the amplifier has a single input and a single output from (16.35), the overall transfer function $w(\mathbf{X})$ of the multiple-loop feedback amplifier becomes

$$w(\mathbf{X}) = D + \mathbf{C}\mathbf{X}\left(\mathbf{1}_p - \mathbf{A}\mathbf{X}\right)^{-1}\mathbf{B} \tag{16.80}$$

When **X** is perturbed to $\mathbf{X}^+ = \mathbf{X} + \boldsymbol{\delta}\mathbf{X}$, the corresponding expression of (16.80) is given by

$$w(\mathbf{X}) + \delta w(\mathbf{X}) = D + \mathbf{C}\left(\mathbf{X} + \boldsymbol{\delta}\mathbf{X}\right)\left(\mathbf{1}_p - \mathbf{A}\mathbf{X} - \mathbf{A}\boldsymbol{\delta}\mathbf{X}\right)^{-1}\mathbf{B} \tag{16.81}$$

or

$$\delta w(\mathbf{X}) = \mathbf{C}\left[\left(\mathbf{X} + \boldsymbol{\delta}\mathbf{X}\right)\left(\mathbf{1}_p - \mathbf{A}\mathbf{X} - \mathbf{A}\boldsymbol{\delta}\mathbf{X}\right)^{-1} - \mathbf{X}\left(\mathbf{1}_p - \mathbf{A}\mathbf{X}\right)^{-1}\right]\mathbf{B} \tag{16.82}$$

As $\boldsymbol{\delta}\mathbf{X}$ approaches zero, we obtain

$$\begin{aligned}
\delta w(\mathbf{X}) &= \mathbf{C}\left[\left(\mathbf{X} + \boldsymbol{\delta}\mathbf{X}\right) - \mathbf{X}\left(\mathbf{1}_p - \mathbf{A}\mathbf{X}\right)^{-1}\left(\mathbf{1}_p - \mathbf{A}\mathbf{X} - \mathbf{A}\boldsymbol{\delta}\mathbf{X}\right)\right]\left(\mathbf{1}_p - \mathbf{A}\mathbf{X} - \mathbf{A}\boldsymbol{\delta}\mathbf{X}\right)^{-1}\mathbf{B} \\
&= \mathbf{C}\left[\boldsymbol{\delta}\mathbf{X} + \mathbf{X}\left(\mathbf{1}_p - \mathbf{A}\mathbf{X}\right)^{-1}\mathbf{A}\boldsymbol{\delta}\mathbf{X}\right]\left(\mathbf{1}_p - \mathbf{A}\mathbf{X} - \mathbf{A}\boldsymbol{\delta}\mathbf{X}\right)^{-1}\mathbf{B} \\
&= \mathbf{C}\left(\mathbf{1}_q - \mathbf{X}\mathbf{A}\right)^{-1}\left(\boldsymbol{\delta}\mathbf{X}\right)\left(\mathbf{1}_p - \mathbf{A}\mathbf{X} - \mathbf{A}\boldsymbol{\delta}\mathbf{X}\right)^{-1}\mathbf{B} \\
&\approx \mathbf{C}\left(\mathbf{1}_q - \mathbf{X}\mathbf{A}\right)^{-1}\left(\boldsymbol{\delta}\mathbf{X}\right)\left(\mathbf{1}_p - \mathbf{A}\mathbf{X}\right)^{-1}\mathbf{B}
\end{aligned} \tag{16.83}$$

where **C** is a row q vector and **B** is a column p vector. Write

$$\mathbf{C} = \begin{bmatrix} c_1 & c_2 & \cdots & c_q \end{bmatrix} \tag{16.84a}$$

$$\mathbf{B}' = \begin{bmatrix} b_1 & b_2 & \cdots & b_p \end{bmatrix} \tag{16.84b}$$

$$\tilde{\mathbf{W}} = \mathbf{X}\left(\mathbf{1}_p - \mathbf{AX}\right)^{-1} = \left(\mathbf{1}_q - \mathbf{XA}\right)^{-1}\mathbf{X} = \left[\tilde{w}_{ij}\right] \tag{16.84c}$$

The increment $\delta w(\mathbf{X})$ can be expressed in terms of the elements of (16.84) and those of **X**. In the case where **X** is diagonal with

$$\mathbf{X} = \operatorname{diag}\begin{bmatrix} x_1 & x_2 & \cdots & x_p \end{bmatrix} \tag{16.85}$$

where $p = q$, the expression for $\delta w(\mathbf{X})$ can be succinctly written as

$$\begin{aligned} \delta w(\mathbf{X}) &= \sum_{i=1}^{p}\sum_{k=1}^{p}\sum_{j=1}^{p} c_i\left(\frac{\tilde{w}_{ik}}{x_k}\right)\left(\delta x_k\right)\left(\frac{\tilde{w}_{kj}}{x_k}\right)b_j \\ &= \sum_{i=1}^{p}\sum_{k=1}^{p}\sum_{j=1}^{p} \frac{c_i\tilde{w}_{ik}\tilde{w}_{kj}b_j}{x_k}\frac{\delta x_k}{x_k} \end{aligned} \tag{16.86}$$

Comparing this with (16.79), we obtain an explicit form for the single-parameter sensitivity function as

$$\mathcal{S}\left(x_k\right) = \sum_{i=1}^{p}\sum_{j=1}^{p} \frac{c_i\tilde{w}_{ik}\tilde{w}_{kj}b_j}{x_k w(\mathbf{X})} \tag{16.87}$$

Thus, knowing (16.84) and (16.85), we can calculate the multiparameter sensitivity function for the scalar transfer function $w(\mathbf{X})$ immediately.

Example 6. Consider again the voltage-series feedback amplifier of Figure 13.9, an equivalent network of which is shown in Figure 16.4. Assume that V_s is the input and V_{25} the output. The transfer function of interest is the amplifier voltage gain V_{25}/V_s. The elements of main concern are the two controlling parameters of the controlled sources. Thus, we let

$$\mathbf{X} = \begin{bmatrix} \tilde{\alpha}_1 & 0 \\ 0 & \tilde{\alpha}_2 \end{bmatrix} = \begin{bmatrix} 0.0455 & 0 \\ 0 & 0.0455 \end{bmatrix} \tag{16.88}$$

From (16.27) we have

$$\mathbf{A} = \begin{bmatrix} -90.782 & 45.391 \\ -942.507 & 0 \end{bmatrix} \tag{16.89a}$$

$$\mathbf{B}' = \begin{bmatrix} 0.91748 & 0 \end{bmatrix} \tag{16.89b}$$

$$\mathbf{C} = \begin{bmatrix} 45.391 & -2372.32 \end{bmatrix} \tag{16.89c}$$

yielding

$$\tilde{\mathbf{W}} = \mathbf{X}\left(\mathbf{1}_2 - \mathbf{AX}\right)^{-1} = 10^{-4}\begin{bmatrix} 4.85600 & 10.02904 \\ -208.245 & 24.91407 \end{bmatrix} \tag{16.90}$$

Also, from (16.13) we have

$$w(\mathbf{X}) = \frac{V_{25}}{V_s} = 45.387 \tag{16.91}$$

To compute the sensitivity functions with respect to $\tilde{\alpha}_1$ and $\tilde{\alpha}_2$, we apply (16.87) and obtain

$$\mathcal{S}(\tilde{\alpha}_1) = \sum_{i=1}^{2}\sum_{j=1}^{2} \frac{c_i \tilde{w}_{i1}\tilde{w}_{1j}b_j}{\tilde{\alpha}_1 w(\mathbf{X})} = \frac{c_1\tilde{w}_{11}\tilde{w}_{11}b_1 + c_1\tilde{w}_{11}\tilde{w}_{12}b_2 + c_2\tilde{w}_{21}\tilde{w}_{11}b_1 + c_2\tilde{w}_{21}\tilde{w}_{12}b_2}{\tilde{\alpha}_1 w} = 0.01066 \tag{16.92a}$$

$$\mathcal{S}(\tilde{\alpha}_2) = \frac{c_1\tilde{w}_{12}\tilde{w}_{21}b_1 + c_1\tilde{w}_{12}\tilde{w}_{22}b_2 + c_2\tilde{w}_{22}\tilde{w}_{21}b_1 + c_2\tilde{w}_{22}\tilde{w}_{22}b_2}{\tilde{\alpha}_2 w} = 0.05426 \tag{16.92b}$$

As a check, we use (14.30) to compute these sensitivities. From (13.45) and (13.52), we have

$$F(\tilde{\alpha}_1) = 93.70 \tag{16.93a}$$

$$F(\tilde{\alpha}_2) = 18.26 \tag{16.93b}$$

$$\hat{F}(\tilde{\alpha}_1) = 103.07 \times 10^3 \tag{16.93c}$$

$$\hat{F}(\tilde{\alpha}_2) = 2018.70 \tag{16.93d}$$

Substituting these in (14.30) the sensitivity functions are:

$$\mathcal{S}(\tilde{\alpha}_1) = \frac{1}{F(\tilde{\alpha}_1)} - \frac{1}{\hat{F}(\tilde{\alpha}_1)} = 0.01066 \tag{16.94a}$$

$$\mathcal{S}(\tilde{\alpha}_2) = \frac{1}{F(\tilde{\alpha}_2)} - \frac{1}{\hat{F}(\tilde{\alpha}_2)} = 0.05427 \tag{16.94b}$$

confirming (16.92).

Suppose that $\tilde{\alpha}_1$ is changed by 4% and $\tilde{\alpha}_2$ by 6%. The fractional change of the voltage gain $w(\mathbf{X})$ is found from (16.79) as

$$\frac{\delta w}{w} \approx \mathcal{S}(\tilde{\alpha}_1)\frac{\delta\tilde{\alpha}_1}{\tilde{\alpha}_1} + \mathcal{S}(\tilde{\alpha}_2)\frac{\delta\tilde{\alpha}_2}{\tilde{\alpha}_2} = 0.003683 \tag{16.95}$$

or 0.37%.

References

[1] F. H. Blecher, "Design principles for single loop transistor feedback amplifiers," *IRE Trans. Circuit Theory,* vol. CT-4, pp. 145–156, 1957.

[2] H. W. Bode, *Network Analysis and Feedback Amplifier Design,* Princeton, NJ: Van Nostrand, 1945.

[3] W.-K. Chen, "Indefinite-admittance matrix formulation of feedback amplifier theory," *IEEE Trans. Circuits Syst.,* vol. CAS-23, pp. 498–505, 1976.

[4] W.-K. Chen, "On second-order cofactors and null return difference in feedback amplifier theory," *Int. J. Circuit Theory Appl.,* vol. 6, pp. 305–312, 1978.

[5] W.-K. Chen, *Active Network and Feedback Amplifier Theory,* New York: McGraw-Hill, 1980, chaps. 2, 4, 5, 7.
[6] W.-K. Chen, *Active Network and Analysis,* Singapore: World Scientific, 1991, chaps. 2, 4, 5, 7.
[7] J. B. Cruz, Jr. and W. R. Perkins, "A new approach to the sensitivity problem in multivariable feedback system design," *IEEE Trans. Autom. Control,* vol. AC-9, pp. 216–223, 1964.
[8] S. S. Haykin, *Active Network Theory*, Reading, MA: Addison-Wesley, 1970.
[9] E. S. Kuh and R. A. Rohrer, *Theory of Linear Active Networks,* San Francisco: Holden-Day, 1967.
[10] I. W. Sandberg, "On the theory of linear multi-loop feedback systems," *Bell Syst. Tech. J.,* vol. 42, pp. 355–382, 1963.